



**South Carolina
Alternate Assessment
(SC-Alt)**

**South Carolina's Alternate Assessment, SC-Alt
Spring 2011 Operational and Field Test Administration**

**Technical Report
October 26, 2011**

**American Institutes for Research
South Carolina Department of Education**

Table of Contents

INTRODUCTION	1
CHAPTER 1: DEVELOPMENT OF ALTERNATE ASSESSMENT IN SOUTH CAROLINA ..3	
<i>Overview of the State Assessment System</i>	3
<i>Purpose of the South Carolina Alternate Assessment</i>	4
<i>Description of the South Carolina Alternate Assessment</i>	4
<i>Background on Alternate Assessment Development in South Carolina</i>	5
<i>Transition from PACT-Alt and HSAP-Alt to SC-Alt</i>	8
<i>American Institutes for Research</i>	8
<i>Development of the Assessment Standards and Measurement Guidelines</i>	9
<i>Stakeholder Input into the Development of the SC-Alt</i>	10
<i>South Carolina State Alternate Assessment Advisory Committee</i>	10
<i>Early Development Activities</i>	10
<i>Summary of the Development and Review of the Original SC-Alt Tasks</i>	11
<i>Content, Bias, and Sensitivity Reviews</i>	12
<i>Development of Field-Test Tasks and Forms</i>	12
<i>Item Data Review</i>	13
<i>Development of Operational Task/Item Pool</i>	13
<i>Design and Development of the 2006–2010 SC-Alt Field Tests</i>	13
<i>Development of the High School Biology Assessment</i>	14
<i>Use of the Student Placement Questionnaires</i>	14
<i>Administration: Placement and Stopping Rules</i>	15
<i>SPQ Summary</i>	15
<i>Teacher Scoring Accuracy</i>	15
<i>2011 Operational Test Booklets and Administration and Scoring Procedures</i>	16
<i>Linking Tasks in Each Grade-Band Assessment</i>	17
CHAPTER 3: SPRING 2011 OPERATIONAL TEST ADMINISTRATION	19
<i>Student Participation for the Spring 2011 Administration</i>	19
<i>Demographics of Participating Students</i>	20
<i>Test Administration Window, Materials, and Timelines</i>	25
<i>Test Administrator Requirements</i>	25
<i>Test Administrator Training</i>	25
<i>Pre-Assessment Using the Student Placement Questionnaire</i>	26
<i>Fidelity of Administration and Accuracy of Scoring</i>	27
CHAPTER 4: SETTING PERFORMANCE STANDARDS	29
<i>Descriptions of Achievement Levels</i>	29
<i>The ID Matching Standard-Setting Process</i>	30

<i>Goals of the Standard Setting</i>	30
<i>Panel Composition</i>	30
<i>Standard-Setting Workshop Activities</i>	30
<i>Cut Score Review and the Setting of Final Cut Scores</i>	31
<i>Consistency of Item Parameter Estimates between the 2010 Biology Field Test and 2011 Operational Assessments</i>	37
<i>Number of Tasks Administered in the Biology Field Test and Operational Administration</i>	38
<i>Scale Score Distributions of the Biology Field Test and Operational Administration</i>	38
<i>Biology Impact Data, Overall and by Demographic Subgroups</i>	40
<i>Impact Data Comparison to 2010 SC-Alt Physical Science Scores</i>	41
CHAPTER 5: TECHNICAL CHARACTERISTICS AND INTERPRETATION OF STUDENT SCORES	43
<i>Analysis and Scaling of Items, Tasks, and Test Forms</i>	43
<i>Assignment of Examinees to Starting Tasks and Item Calibration and Test Forms Linking</i>	43
<i>Analysis Plan</i>	44
<i>Data Preparation and Quality Control</i>	44
<i>Classical Item Analysis</i>	45
<i>Review of Items Not Meeting the Specified Psychometric Criteria</i>	46
<i>Item Response Theory Calibration and Linking Test Forms</i>	46
<i>Using Item Responses to Estimate Student Proficiency</i>	50
<i>Test Score Reliability</i>	51
<i>Classification Accuracy</i>	54
<i>Calculation of the Probability of Being Classified Above a Cut Score Given the Student's Theta Score</i>	55
<i>Calculation of the Likelihood of Theta Given Item Scores z and Step Parameters b</i>	56
CHAPTER 6: SCORE REPORTS	57
CHAPTER 7: STUDENT PERFORMANCE DATA FROM THE SPRING 2011 ADMINISTRATION	59
CHAPTER 8: VALIDITY	61
<i>Content Validity</i>	61
<i>Convergent and Discriminant Validity</i>	63
<i>Selection of Traits and Methods</i>	64
<i>Results</i>	64
<i>Validity of the Student Placement Questionnaire (SPQ)</i>	66
<i>Number of Tasks Administered</i>	71
<i>Number of Items Administered</i>	77
<i>Achievement Level of Students by Start Point</i>	79
<i>Summary</i>	79

REFERENCES	81
------------------	----

APPENDIX A: ASSIGNMENT OF TASKS TO GRADE-BAND FORMS FOR THE SPRING 2011 ADMINISTRATION

APPENDIX B: STARTING AND STOPPING RULES FOR USING THE STUDENT PLACEMENT QUESTIONNAIRE

APPENDIX C: SCORING AUDITS AND ANALYSIS OF VIDEO-RATER DATA FROM THE SPRING 2011 OPERATIONAL ADMINISTRATION

APPENDIX D: DESCRIPTIONS OF ACHIEVEMENT LEVELS (DALS)

APPENDIX E: SUMMARY OF LINKING DESIGN

APPENDIX F: STATISTICS SUMMARIES FOR THE 2011 SPRING FIELD-TEST ITEMS

APPENDIX G: MARGINAL RELIABILITY BY GRADE-BAND, SUBJECT, STARTING TASK, GENDER, AND ETHNIC GROUP

APPENDIX H: SCORE REPORT SAMPLE

APPENDIX I: STUDENT PERFORMANCE BY DEMOGRAPHICS, GRADE-BAND, AND SUBJECT AREA

Introduction

This report details the design, development, and spring 2011 operational and field test results for the South Carolina Alternate Assessment (SC-Alt). The SC-Alt consists of five content areas: English language arts (ELA), mathematics, science, social studies, and (high school) biology. The assessments are administered across three grade-bands: 3–5, 6–8, and 10. The new high school biology assessment was administered operationally for the first time in spring 2011.

Chapter 1: *Development of Alternate Assessment in South Carolina* describes the background of the alternate assessments in South Carolina, the format of the previous assessments, and the need for a new alternate assessment.

Chapter 2: *Test Development* describes the design of the alternate assessment and the development of tasks and items to measure academic growth among students who have significant cognitive disabilities. The Student Placement Questionnaire (SPQ), a unique feature designed to maximize the efficiency of teacher and student testing time, is described and thoroughly reviewed. The development of a vertical scale linking grade-appropriate tasks across grade levels and complexity levels within grades is described.

Chapter 3: *Spring 2011 Operational Test Administration* details the spring 2011 operational test administration in ELA, mathematics, science, social studies, and (high school) biology; test administrator training; use of the SPQ; measures taken to ensure the accuracy of scoring; and the maintenance of test security.

Chapter 4: *Setting Performance Standards* describes the procedures for setting performance standards. The chapter includes a summary of the Item Descriptor (ID) Matching procedure, the goals of the standard-setting workshops, the composition of the standard-setting panels, the workshop activities, and the panels' recommended performance standards. This chapter also presents an impact analysis of the biology standards, based on the data from the spring 2011 operational administration.

Chapter 5: *Technical Characteristics and Interpretation of Student Scores* reviews technical topics including analysis and scaling, reliability of test scores, the procedures used to calculate internal consistency reliability estimates, and classification accuracy estimates.

Chapter 6: *Score Reports* describes the score reporting system for SC-Alt with emphasis on the Individual Student (Family) Report (see Appendix H) from which the summary reports are derived, the information contained in the various reports, and their intended uses.

Chapter 7: *Student Performance Data from the Spring 2011 Administration* provides an overview of statewide achievement on the SC-Alt, based on the spring 2011 operational test administration.

Chapter 8: *Validity* reports on content validity and convergent and discriminant validity topics as well as the validity of the SPQ and the operational performance of the tailored assessment under the SPQ's start/stop rules.

Chapter 1: Development of Alternate Assessment in South Carolina

Overview of the State Assessment System

The South Carolina Assessment System includes the South Carolina Palmetto Assessment of State Standards (PASS), the High School Assessment Program (HSAP), and the End-of-Course Examination Program (EOCEP). These state-level assessments are required by the Education Accountability Act of 1998 (EAA) as amended May 2008 and are aligned with the state's academic standards for each subject and grade level.

- PASS measures the performance of all public school students in grades 3 through 8 in the content areas of English language arts (ELA), mathematics, science, and social studies.
- HSAP measures the performance of high school students in ELA and mathematics and is used both as one criterion for eligibility to receive a high school diploma and as the primary source for reporting the federally mandated data required by the No Child Left Behind Act (NCLB).
- EOCEP is administered in gateway courses at the high school level. The Biology EOCEP examination is counted for participation purposes for NCLB reporting.

The EAA establishes a performance-based accountability system that includes all students. This act supports South Carolina's commitment to public education and a conviction that high expectations for all students are a vital component of improving academic education.

The goals of the state assessment system are as follows:

- Increasing academic performance of all children and, ultimately, raising high school graduation rates
- Implementing rigorous academic achievement standards that are aligned with the South Carolina curriculum standards
- Improving instruction based, in part, on the implementation of these higher standards
- Using the results of challenging assessments that measure student performance relative to these standards

Another goal is to inform various audiences—teachers, school administrators, district administrators, South Carolina State Department of Education (SCDE) staff, parents, and the public—of the status of academic performance and of the progress of public school students toward meeting South Carolina's academic achievement standards.

The South Carolina academic standards form the basis for alignment across the state education system for district and school curricula, classroom instruction, units of study, and learning experiences. **The academic standards are the basis for all assessments in the state assessment system, including the alternate assessment.**

Purpose of the South Carolina Alternate Assessment

The purpose of the alternate assessment based on alternate achievement standards is to capture and evaluate the performance of students who have traditionally been excluded from statewide testing programs and to improve instruction for these students by promoting appropriately high expectations and the inclusion of these students in state accountability for district report cards and for adequate yearly progress (AYP) reporting at the school, district, and state levels.

Description of the South Carolina Alternate Assessment

The South Carolina Alternate Assessment (SC-Alt) is administered to students who have been determined by the Individualized Education Program (IEP) team to be unable to participate in the general state assessments even with appropriate accommodations. It is an alternate assessment on alternate achievement standards to the PASS for students in grades 3–8 and the HSAP and Biology EOCEP for high school students.

The test is administered to students who meet the participation criteria for alternate assessment and who are of the ages of typical students in grades 3–8 and 10. Students who are ages 8–13 (the typical ages for grades 3–8) are assessed in ELA, mathematics, science, and social studies. Students who are 15 (the typical age of students in grade 10) are assessed in ELA, mathematics, and biology.

The SC-Alt consists of a series of performance tasks that are scored by the test administrator (teacher) as they are administered. The performance tasks are scripted activities, and each task contains four to eight related items. The items have a scaffolded scoring script to reduce the complexity of the item when students do not respond successfully on the first attempt. All items are linked to the South Carolina academic content standards through the South Carolina Alternate Assessment Extended Standards. The Extended Standards are linked explicitly to the South Carolina academic standards for grades 3–8 and 10, although at less-complex or prerequisite levels. The SC-Alt has three forms: elementary, middle, and high school. Students' assignment to forms is based on their age on September 1 of the tested year; 8- to 10-year-olds take the elementary form, 11- 13-year-olds take the middle school form, and 15-year-olds take the high school form.

The assessment is designed to minimize the teacher and student testing burden by administering only those items that are well-suited to a student's achievement level. The test administrator completes a Student Placement Questionnaire (SPQ) to determine the most appropriate starting task for the student. Tasks are arranged in ascending order of difficulty. Once the appropriate starting task is identified, test administrators continue to administer tasks until the student can no longer respond successfully.

The first operational administration of the SC-Alt was conducted during a seven-week testing window during spring 2007 in ELA, mathematics, and science. A census field test was conducted during the same assessment window for the social studies assessment. In spring 2009 and spring 2011, embedded field tests in ELA, mathematics, science, and social studies were added. Also in 2011, high school biology was introduced as an operational assessment. Documentation related to the 2011 operational administration is the focus of this Technical Report.

Background on Alternate Assessment Development in South Carolina

The 1997 amendments to the Individuals with Disabilities Education Act (IDEA '97) created the mandate to include all children, including children with significant disabilities, in state testing and accountability systems. The vision for the South Carolina alternate assessment system was initiated in early 1998 in response to the IDEA '97 regulations. This vision has driven the development and revision of alternate assessment in South Carolina.

A core team of staff from the SCDE Offices of Exceptional Children, Assessment, Research, and Curriculum and Standards met in March 1998 to develop a plan for designing an alternate assessment to meet the IDEA mandate and to be included in the state assessment system. The team's first steps were to convene a steering committee and seek technical assistance from the Mid-South Regional Resource Center (MSRRC) to explore strategies for designing an alternate assessment.

The Alternate Assessment Steering Committee was convened on May 12, 1998, to assist SCDE in determining how to include students with significant cognitive disabilities in statewide assessments. The committee comprised parents, special education and general education teachers, administrators, and representatives from other agencies. Dr. Ken Olsen of MSRRC provided the committee with technical assistance, including information on IDEA requirements, examples of options that some states were using or considering, and research available on alternate assessment. He facilitated a process that allowed the Steering Committee to reach shared foundational beliefs, address eligibility criteria and content and performance standards, and develop plans.

To ensure that all students, including students with significant disabilities, are included in the testing and accountability systems and have appropriate access to instruction in the South Carolina academic standards, the Steering Committee determined that the alternate assessment would be based on the following principles:

- All children can learn, be expected to meet, and be challenged to meet high standards.
- Special education is an extension and adaptation of the general education program and curriculum, rather than an alternate or separate system.
- The South Carolina State Board-approved standards are the foundation for all students, including students with unique needs and abilities.
- Measurement and reporting must be defensible in terms of feasibility, validity, reliability, and comparability.
- Results of the state standards-based program must be used to improve planning, instruction, and learning.
- An alternate assessment is appropriate for the few students for whom the state assessment, even with accommodations, is not appropriate.
- The alternate assessment is designed for a diverse group of students and should be flexible enough to address their individual needs.

The committee articulated these goals for the alternate assessment:

- To provide evidence that students have acquired the skills and knowledge necessary to become as independent as possible
- To document the student's performance and the performance of the programs serving the student
- To merge instructional best practice, instruction in state standards, and assessment activities
- To provide information in the development of curriculum that is responsive to the student's needs

The Steering Committee created the following participation guidelines to guide IEP team decisions regarding students who should participate in the alternate assessment:

- The student demonstrates significant cognitive disabilities and adaptive skills, which result in performance that is substantially below grade-level achievement expectations even with the use of accommodations and modifications.
- The student accesses the state-approved curriculum standards at less complex levels and with extensively modified instruction.
- The student has current adaptive skills requiring extensive direct instruction and practice in multiple settings to accomplish the application and transfer of skills necessary for application in school, work, home, and community environments.
- The student is unable to apply or use academic skills across natural settings when instructed solely or primarily through classroom instruction.
- The student's inability to achieve the state grade-level achievement expectations is not the result of excessive or extended absences or social, cultural, or economic differences.

NOTE: The term significant cognitive disabilities was added by the South Carolina Alternate Assessment Advisory Committee to the criteria after the passage of the NCLB December 2003 regulations on alternate assessment.

The Steering Committee recommended that the state develop a portfolio collection of evidence of student progress toward the South Carolina academic standards similar in design to the Kentucky Portfolio Alternate Assessment. The committee also recommended that SCDE prepare a Request for Proposal (RFP) for a contractor to develop the alternate assessment. Advanced Systems in Measurement and Evaluation Inc. (ASME), which later became Measured Progress, was awarded the contract. This company, along with the Inclusive Large Scale Standards and Assessment (ILSSA) project at the University of Kentucky, began work with SCDE on the design of PACT-Alt.

A work group was convened to define the domain for instruction and assessment. To ensure that the South Carolina curriculum standards were the foundation for all students, including students with unique needs and abilities, the work group developed adaptations of the curriculum standards. The work group comprised special education teachers, regular education teachers,

parents, administrators, higher education personnel, representatives from community agencies, and SCDE personnel. The work group process, which was facilitated by staff from MSRRRC, focused on the prerequisite skills found primarily in the curriculum standards in prekindergarten through grade 2.

The work group affirmed that special education services must operate as an extension of the general education program and curriculum rather than as an alternate or separate system. The standards in this initial document were identified as concepts that every student, including students with moderate to severe disabilities, should know or be able to perform. These selected standards, which focused on skills that were deemed essential and attainable for every student, were directed toward the following goals:

- Enhancing the quality of students' communication skills
- Improving the quality of students' everyday living
- Improving students' ability to function in society and promoting in them an acceptance of and respect for self and others
- Preparing students for transition into adult living
- Moving students toward independence, which may range from a level of self-care with assistance to total self-sufficiency

The extensions were based on the state academic content standards in prekindergarten through grade 2. For each selected standard, examples of essential real-world performance skills were developed. The articulation of these performance skills was designed to provide the rationale for teaching the standards and to serve as guides for teachers and parents regarding how the student demonstrated a skill. The committee specified that these performance skills could be accomplished in home, school, and community environments through a variety of individualized communication systems and might incorporate a variety of supports, such as physical assistance, physical prompts, verbal prompts, and technology. The document, *The Extensions and Adaptations of the South Carolina Curriculum Standards for Students Participating in Alternate Assessment*, became the focus of the portfolio assessment process, HSAP-Alt performance tasks, and professional development training. In 2002, this document was revised and renamed the *Resource Guide to the South Carolina Curriculum Standards for Students in Alternate Assessment*, but it was still aligned to curriculum standards for prekindergarten through grade 2. This work was based on the IDEA requirements and the thinking at the time about how students with significant cognitive disabilities should be included in the general education curriculum and assessment.

Beginning with the 2000–2001 school year, students in grades 3–8 who met the participation criteria for alternate assessment were assessed with the portfolio assessment, PACT-Alt. In 2003, a high school assessment, HSAP, which was designed to meet AYP requirements, was added to the state assessment system, and an alternate to HSAP was developed to measure student proficiency in ELA and mathematics. A Stakeholder Committee with expertise in high school instruction of students with significant cognitive disabilities and academic standards was convened to guide the development of the high school alternate assessment, HSAP-Alt. The committee recommended designing an assessment based on performance on a series of tasks

linked to the state curriculum standards. The HSAP-Alt consisted of a series of scripted performance tasks in ELA and mathematics with scaffolded administration and scoring procedures aligned with the *Resource Guide to the South Carolina Curriculum Standards for Students in Alternate Assessment*.

One critical piece of the development and implementation process of PACT-Alt and HSAP-Alt was the provision of intensive professional development related to standards-based instruction, much of it based on the work of Harold Kleinberg and Jacqui Farmer Kearns. A resource for professional development was their book, *Alternate Assessment: Measuring Outcomes and Supports for Students with Disabilities*. Professional development was essential to the implementation of the portfolio assessment because the teacher was responsible for teaching the student the content related to the academic standards, assessing the student's progress, and providing evidence of the instruction and progress in the portfolio. Prior to the implementation of the alternate assessment and the IDEA requirement to include students with disabilities in the general education curriculum, many students with disabilities, especially those with significant disabilities, and their teachers had been excluded from standards-based instruction and professional development related to academic standards.

Transition from PACT-Alt and HSAP-Alt to SC-Alt

After seeking input on the vision of a new alternate assessment on alternate achievement standards from the Advisory Committee and teachers who were conducting alternate assessment, SCDE wrote an RFP for the redesign or design of the alternate assessment system. The design was to be consistent with South Carolina's commitment to the instruction and assessment of students with significant cognitive disabilities and NCLB requirements. The focus was to be on grade-level academic standards. The new system was to address concerns related to teacher burden and time involved in assessment while supporting improved instruction based on state academic achievement standards. Extensive training for test administrators was to be integrated into the design of the assessment.

In September 2004, a contract was awarded to American Institutes for Research (AIR) to assist the state in revising the alternate assessment. AIR managed the administration and analyses of the PACT-Alt and HSAP-Alt assessments during the 2004–2005 and 2005–2006 school years while developing the new alternate assessment, the South Carolina Alternate Assessment (SC-Alt), with SCDE.

American Institutes for Research

American Institutes for Research (AIR) has more than 50 years of experience as a nonprofit organization dedicated to assessment, behavioral science, and educational research. AIR developed the South Carolina HSAP and the EOCEP programs and has enjoyed a successful collaboration with SCDE for a number of years.

Chapter 2: Test Development

The South Carolina academic content standards are the basis for alignment across the state for district and school curricula, classroom instruction, units of study, and learning experiences. The curriculum standards are the basis for the Palmetto Assessment of State Standards (PASS), the High School Assessment Program (HSAP), the End-of -Course Examination Program (EOCEP), and the alternate assessment. An initial step in the design of the alternate assessment was developing Assessment Standards and Measurement Guidelines (ASMGs).

Development of the Assessment Standards and Measurement Guidelines

In April 2005, a committee comprising South Carolina special education teachers, content specialists, SCDE staff, and AIR staff designed the ASMG document to support the new assessment development. The process involved extending the state academic standards in ELA, mathematics, science, and social studies in grade-bands 3–5, 6–8, and 10 to be accessible to students with significant cognitive disabilities. This document replaced the *Resource Guide to the South Carolina Curriculum Standards for Students in Alternate Assessment*.

The ASMGs were the foundation for the development of the assessment tasks for the SC-Alt. The ASMGs in each content area are distillations of the essence of South Carolina curriculum standards at each grade level.

Each content area committee reviewed the large array of standards and prioritized those in grade-bands 3–5, 6–8, and 10 that they deemed most important to students now and in the future. They then reduced the complexity of these standards, while retaining the essence of the grade-level content knowledge and skills, to make the academic standards appropriate and accessible for students with significant cognitive disabilities. The committee was careful to address both the depth and the breadth of the academic standards and used professional judgment based on experience with the population and the content to determine the standards to be assessed. The resulting document provided the link to the grade-level standards and indicators in the state academic standards. The measurement guidelines gave task writers and teachers the specificity necessary to translate the assessment standards into assessment tasks and items and classroom instruction. A list of individuals who were involved in this process is included in each ASMG content document.

NOTE: The ELA committee recommended that the standards in the Research Goal not be included in the assessment standards. The rationale for this recommendation was that this goal was not tested to any great extent in PACT because this content is primarily taught and assessed at the classroom level. Committee members, however, indicated that the Communication Goal included standards that they deemed very important to this population, and they recommended including assessment standards for this strand.

The State Board of Education adopted revised mathematics and ELA academic standards in August 2007 and May 2008. The State Board of Education required replacement of the high school physical science end-of-course assessment for all students with a biology end-of-course assessment. The adoption of these revised standards, which occurred outside the cyclical review timetable, and the replacement of the physical science end-of-course assessment with the biology end-of-course assessment had a direct impact on the ongoing schedule for developing additional tasks for the task pool.

During the 2007 and 2008 school years, committees of special educators and general educators met to extend the revised ELA, mathematics, and science academic standards, as well as the biology standards. These documents were designed to provide specificity for instruction as well as assessment, so the committees extended all standards and indicators including those for non-tested grades. These documents, referred to as the Extended Standards, replaced the ASMGs in ELA, mathematics and science, and provided extensions for biology. The Extended Standards provide extensions for all grade levels, including those that are not tested, and guidance to assist educators with instructional access to the state academic standards.

Stakeholder Input into the Development of the SC-Alt

To ensure the validity of the overall assessment process, a great deal of time and effort was spent obtaining input from various sources, including the State Alternate Assessment Advisory Committee, classroom teachers, parents, and other agency personnel.

South Carolina State Alternate Assessment Advisory Committee

The State Alternate Assessment Advisory Committee meets to provide oversight to the SC-Alt. The committee includes members of the original Alternate Assessment Steering Committee and the High School Stakeholder Committee. The committee also includes parents, special educators, and representatives of higher education, content specialists, special education directors, and district test coordinators. Additional members include representatives from the Department of Disabilities and Special Needs, the University of South Carolina School of Medicine, the South Carolina Assistive Technology Project, the South Carolina Interagency Deaf-Blind Project, the Autism Society of South Carolina, and Pro-Parents of South Carolina.

The Advisory Committee provided input on its expectations for the revised alternate assessment during the first meeting with the contractor, AIR, on November 5, 2004. SCDE and AIR staff reported each step of the development process to the Advisory Committee at each meeting and sought its advice and recommendations.

Early Development Activities

At the recommendation of the Advisory Committee, AIR item writers visited classrooms in South Carolina during January and February 2005 to observe teaching strategies and materials that were in use. They also reviewed PACT-Alt portfolios for examples of evidence that teachers used to demonstrate progress toward proficiency on grade-level standards and examined the characteristics of the HSAP-Alt performance event in order to build on the existing system.

Teacher focus groups convened during January 2005 obtained feedback from teachers on the types of tasks they believed were appropriate, the protocol format they preferred, and the materials they recommended for inclusion in the assessment.

Qualified item writers employed by AIR were trained to write tasks and items specifically aligned with the ASMGs. Item writing teams included AIR staff with expertise in the content areas; alternate assessment specialists; and consultants in the areas of instruction of students who are blind and visually impaired, students who are deaf and hard of hearing, and students with cognitive disabilities.

On February 14, 2006, prior to the development of science and social studies tasks, SCDE staff and the AIR alternate assessment specialist provided additional training to the writing teams. The training was based on *Designing from the Ground Floor*, materials developed by the National Alternate Assessment Center (2005).

Consideration of universal design was a focus throughout the development process. Items, including passages and response options, were developed to use objects, pictures, picture symbols, words, and numbers. Several tasks in all four content areas and at different levels of complexity were piloted with South Carolina teachers and students in March and May 2005. AIR staff then interviewed the pilot teachers to determine the item characteristics and parameters that teachers believed worked well or did not work.

Summary of the Development and Review of the Original SC-Alt Tasks

- The task and item development process began with the creation of task kernels. AIR was primarily responsible for the majority of task kernels, with input from SCDE and teachers in South Carolina. Tasks kernels are basic ideas for an assessment activity, stimulus materials, and purpose, which, based on their relation to the South Carolina ASMGs, were used to develop a task and its items.
- SCDE reviewed the task kernels and provided feedback to AIR on which kernels were acceptable, which were unacceptable, and which needed revision. These reviews included alignment with the ASMGs.
- AIR item writers developed the items and stimulus materials. These items were reviewed internally by the content experts for clarity, quality, and alignment with the ASMGs.
- Following the comprehensive AIR internal review, the tasks and items underwent technical review by AIR to ensure that the items were properly keyed and scaffolded, the instructions were appropriate, the stimulus materials were interpretable, and the items were generally consistent in design with other tasks and items under development.
- Items that passed internal review by the AIR development staff were reviewed by the senior content lead for each content area and the senior alternate assessment specialist. This review ensured that within the content area, tasks and items followed the design of the assessment and were consistent with respect to format, presentation, and general administration procedures.
- Before items were passed to SCDE, the project director reviewed all items to ensure that they were consistent with the foregoing factors across content areas and grade bands.
- Following the final internal AIR review, items were passed to SCDE for its review. During this process, SCDE staff, including content specialists, special educators, and assessment specialists, provided feedback to AIR on the design of the tasks and items, the alignment of items to the ASMGs, and the appropriateness of the items for use in South Carolina. Some items were revised by SCDE to improve alignment with the ASMGs.
- Approved items were placed into tasks for a small-scale tryout, conducted by AIR with the assistance of teachers in South Carolina and Northern Virginia and AIR staff. These tryouts provided invaluable information regarding the clarity of instructions, the utility of

the stimulus materials, and the success of the items and tasks in producing expected responses. Items that showed obvious problems were revised or discarded.

- After changes were made to the prototypes as a result of the pilots and tryouts, a committee of South Carolina teachers was convened on July 12, 2005, to review the revised tasks and provide further input and recommendations.

Content, Bias, and Sensitivity Reviews

Once small-scale tryouts were concluded, AIR, SCDE, and educators in South Carolina reviewed the tasks and items for alignment with the ASMGs and for bias and sensitivity concerns. The reviews for content and bias and sensitivity were combined due to the direct impact of the task format, materials, and language on the assessment accessibility for the population. Committees comprising teachers of students with significant cognitive disabilities, representatives of higher education, special education administrators, experts in the instruction of students with limited English proficiency, and content experts from across the state participated in these reviews to consider the following:

- Alignment to the ASMGs and Extended Standards
- Bias for specific groups and types of disabilities
- Accessibility of the tasks to the entire population for whom the test was designed
- Characteristics that might lead to bias or are inappropriate for or insensitive to the nature of the student subgroups (e.g., exclusionary language, stereotypes)
- Format and content of the tasks
- Accessibility of materials
- Clarity of instructions and ease of administration

The review committee meetings were conducted in November 2005, May 2006, and, for the spring 2009 embedded field test, in November 2008. For the 2010 biology field test, the content and bias and sensitivity review meetings were held in June 2009. The committee reconvened in July 2010 to review newly developed field-test tasks for the 2011 administration. During the reviews, committee members recommended that some items be revised or eliminated.

Development of Field-Test Tasks and Forms

- On the basis of the feedback from all the steps above, AIR conducted a final review and sign-off for all items and tasks. Following this review, the items and tasks were affirmed ready for field-testing.
- Prior to assembling tasks into test forms, the senior content lead for each content area and the project director reviewed the items and tasks one last time to determine whether the revisions were appropriate and maintained the alignment of the item to the targeted standard.

- For stand-alone field tests, tasks and their items were then placed into field-test forms consistent with the specifications described earlier. For embedded field tests, the tasks and their items were placed into designated locations on the operational test forms.

Item Data Review

- After field-testing, AIR and SCDE staff, including alternate assessment specialists, psychometricians, content specialists, and special educators, met to review the field-test statistics.
- They reviewed the statistics associated with each item and task to determine whether the items were functioning within expectations and whether the tasks were appropriately placed within the instrument. The statistical criteria applied to the field-test item data and to the operational item data are described in Chapter 5.
- The committee also considered teacher comments on specific items from the field test, data from field-test observations, and the results of the alignment studies to make decisions about the inclusion of items in the operational assessment.
- Items that did not meet these criteria were retained for possible future operational use or were revised for recalibration.
- The Item Data Review meetings for the original independent field tests were conducted in August 2006 and June 2007. The Item Data Review of the 2010 independent biology field test was held in July 2010. The other administrations after 2007 used an embedded field-testing approach. For the embedded field tests, item data reviews were conducted in 2008 for social studies and in 2009 and 2011 for ELA, mathematics, science, and social studies.

Development of Operational Task/Item Pool

- AIR once again reviewed all data associated with the tasks and items to determine whether the items were functioning as expected and were useful for measuring the achievement of students in South Carolina.
- Items that survived all review and analysis criteria were placed into the operational task/item pool.

Design and Development of the 2006–2010 SC-Alt Field Tests

Following the task development process, the field-test forms were designed and produced. The primary purposes of the independent field-test administrations for English language arts and mathematics (spring 2006), science (fall 2006), and social studies (spring 2007) were to produce data to evaluate SC-Alt tasks and items and to guide the assembly of operational test forms to be used in 2007 and beyond. Student scores based on field-test data were not reported.

An embedded field test (spring 2008) tested the symbolate version of the social studies task “George Washington” so that its performance could be compared with the text version used in the spring 2007 field test.

The design, data collection, and analysis of the independent 2006 and 2007 field tests in ELA, mathematics, science, and social studies, of the 2008 embedded social studies field-test tasks, and of the 2009 and 2010 embedded field-test tasks in ELA, mathematics, science, and social studies were discussed in the spring 2007, 2008, 2009, and 2010 operational technical reports.

Development of the High School Biology Assessment

During spring 2010, concurrently with the operational SC-Alt administration, 21 new high school biology tasks were field-tested on 15- and 16-year-old students eligible for alternate assessment. The biology field test was administered to 472 students on two forms of 12 tasks each. The forms were linked by three shared tasks, which allowed all biology items to be calibrated on the same scale.

The IRT parameters, classical item statistics, and fit and DIF statistics were subjected to an item data review conducted with AIR and SCDE staffs on July 20, 2010. A standard-setting workshop based on the biology field-test data was conducted on September 14 and 15, 2010. Biology was subsequently administered operationally for the first time in the spring 2011 assessment.

Use of the Student Placement Questionnaires

The Student Placement Questionnaires (SPQs) are brief structured rating instruments that represent the range of communication levels and cognitive-academic functioning found in the population of alternate assessment examinees. AIR developed the SPQ for the South Carolina Alternate Assessment program.

The student placement process is intended to achieve several important goals:

- It matches student achievement levels with the difficulty of the tasks and items that are administered.
- It allows a maximum number of student item responses at an appropriate level of difficulty.
- It minimizes fatigue by targeting the assessment to the student.
- It supports the psychometric rigor of student scores. A student is administered a better targeted test than one that contains many items the student might find too difficult. Better test targeting contributes to better score reliability. Because fatigue effects from the student's limited attention span are reduced, the validity of the overall assessment is enhanced.

Teachers completed the SPQs in each content area to identify the most appropriate starting task for each student. For each subject, the SPQs prompted the teacher with between 12 and 15 "can do" questions (e.g., can this student recognize the sun, moon, Earth?). The questions were grouped by major content standards and sampled across low-, moderate-, and high-complexity levels. Each question rated the student's functioning on a 4-point scale, valued 0 to 3. Answering the 12 to 15 questions of each SPQ, summing the total score, and identifying the most appropriate starting task in a lookup table took test administrators approximately 6 or 7 minutes.

The lookup table identified ranges of SPQ scores that corresponded to one of three starting tasks. Teachers used the SPQs to assign students to starting points on the assessment. Cut points for the

science SPQ were based on the rules derived for the mathematics SPQ but were altered for the number of items on the science SPQ. Details regarding the student participation, analysis, and conclusions drawn from use of the SPQ placement procedure appear below.

Administration: Placement and Stopping Rules

After teachers identified the most appropriate starting task for a student, they followed several rules as they administered the starting task and subsequent tasks. If starting at task 1, the teacher would administer at least six operational tasks; otherwise, at least seven operational tasks would be administered. For detailed placement and stopping rules for the spring 2011 operational and field-test administrations, see Appendix B.

SPQ Summary

The preceding discussion reviewed some of the implementation procedures for the SPQ. Here we review two of the technical characteristics of the SPQ: the method used to select the SPQ recommended starting task and the usefulness of the SPQ as an indicator of student starting task.

The technical development of the SPQ and determination of the cut points to determine starting tasks are fully described in American Institutes for Research, 2008, *South Carolina Alternate Assessment (SC-Alt): Technical Report for English Language Arts and Mathematics Field Test Administration, Spring 2006*.

Usefulness of the SPQ for Determining the Starting Task. AIR has gathered information regarding the agreement between the SPQ recommended start points and the final observed start points by reviewing item data following each operational administration. The results of the study of 2011 data are reported in detail in Chapter 8.

Use of the SPQ pre-assessment score is only the first step in the procedure used by the test administrator in determining where the student should start the assessment. The instructions for using the SPQ include procedures requiring teachers to adjust the starting point below the SPQ recommended start point when the student is not successful on the first administered task. Alternately, after reviewing the assessment, some teachers may have judged that a student needed to start at a higher level than recommended by the SPQ.

The results of the 2011 study indicate that the agreement between the SPQ recommended start point and the observed start point by content area were 98% for ELA, 98% for mathematics, 99% for science/biology, and 98% for social studies. Since the test administrator is required to make adjustments based on the student's success on the first task, and these adjustments are reflected in the agreement rates, the SPQ appears to be working very effectively for targeting the first task to begin the assessment process.

The results of the Start-Stop Analysis reported in Chapter 8 also support the effectiveness and validity of the SPQ and the SC-Alt tailored assessment design.

Teacher Scoring Accuracy

The design of the SC-Alt includes test administrator (teacher) scoring of student responses. The degree of accuracy with which the test administrator evaluates student performance determines whether the student receives the correct scores and the correct performance level.

A video study and a second rater study were conducted during the 2011 administration to confirm that test administrators were following all scoring procedures accurately. For these studies, scoring accuracy refers to the degree to which teachers follow scaffolding and scoring directions correctly and assign correct scores to student responses. In the video study, scoring accuracy by the test administrators was evaluated by having trained raters at AIR review the videotapes of the test administrations and score the student responses without knowledge of the scores assigned by the test administrators. A pilot sample of elementary school students had their item responses scored simultaneously by a second rater who was present during the test administration. This pilot sample was not videotaped. After the raters concluded their scoring of the student responses, the consistency between the test administrators and AIR raters was determined.

Detailed results of the scoring consistency analysis are presented in Appendix C. The results indicated that there was a high degree of consistency between the scoring of the test administrators and the AIR raters, suggesting that test administrators in South Carolina understood the scoring procedures and implemented them accurately when scoring student responses. The two studies yielded comparable results.

2011 Operational Test Booklets and Administration and Scoring Procedures

For each grade-band test form in each content area, tasks and items were selected that met the statistical criteria and that covered the breadth of the targeted Extended Standards. The 2011 operational test forms in ELA, mathematics, science, and social studies were revised by inserting embedded field-test tasks in each grade-band form. In addition, two operational field-test tasks were included in each of the science grade band forms. High school biology was first administered operationally in spring 2011. All operational forms had their tasks ordered by increasing difficulty of the items in each task, which was determined by Item Response Theory (IRT) analysis. The goal was to use technically sound assessment instruments to support valid inferences about what students know and can do relative to the Extended Standards in each content area.

The SC-Alt operational administration in spring 2011 included three sets of test materials in English language arts, mathematics, and science/biology: one for the 3–5 grade-band assessment, one for the 6–8 grade-band assessment, and one for the grade 10 assessment. The social studies assessment used two sets of materials, one each for grade-bands 3–5 and 6–8 (grade 10 is not part of the social studies assessment). Similarly, science was administered only in grade-bands 3–5 and 6–8, while the grade 10 science assessment of previous administrations was replaced by biology. Teachers (test administrators) received a *Test Administration Manual (TAM)* and comprehensive training based on the manual and the test materials.

The 2011 test booklets for English language arts, mathematics, and social studies contained 12 operational tasks and 3 embedded field-test tasks. The elementary and middle school science assessments had 10 operational tasks, two operational field-test tasks, and three embedded field-test tasks.¹ Finally, the biology assessment contained 12 operational tasks and no embedded

¹ The items of operational field-test tasks were field-tested with the operational sample, calibrated to the operational item bank, subjected to an item data review, and then employed to score the participating students operationally. The items of embedded field-test tasks were also calibrated on the operational sample and subjected to an item data review but were not used for operational student scoring during the current administration.

field-test tasks. Operational tasks were arranged in test forms in the order of the empirical difficulty of the items in each task. The ELA, mathematics, and social studies test forms (elementary, middle, and high school) included linking tasks to support psychometric linking of the grade-band score scales. Each task consisted of four to eight separate items. Teachers were instructed to administer a minimum of six or seven operational tasks to each student, depending on the SPQ-designated starting point, and to continue administration of subsequent tasks until the student was no longer successful.

Teachers also received other materials with each test booklet:

- Physical manipulatives
- Printed manipulatives
- An answer folder for each participating student
- A Student Placement Questionnaire and directions for determining the starting task for each student (included in the answer folder)

Exhibit 2.1 summarizes the operational grade-band assessments and the numbers of operational tasks in each grade assessment for 2011.

Exhibit 2.1: Numbers of Operational and Field-Test Tasks in Each Grade-Band Assessment, 2011

Grade-Band	Total in Each Grade-Band (Field-Test Tasks in Parentheses)				
	ELA	Math	Science	Social Studies	Biology
3–5	12 (+3 FT)	12 (+3 FT)	10 + 2 OFT (+3 FT)	12 (+3 FT)	
6–8	12 (+ 3 FT)	12 (+3 FT)	10 + 2 OFT (+3 FT)	12 (+3 FT)	
10	12 (+3 FT)	12 (+3 FT)			12

Note: FT = embedded field test task; OFT = operational field test task.

The approximate test length for each grade-band assessment for the 2011 administration was 60 items (12 tasks × an average 5 items per task) and 120 score points (60 items × an average 2 points per item).

Linking Tasks in Each Grade-Band Assessment

All tasks in each SC-Alt grade-band assessment are aligned to the extended standards in that grade-band. Because adjacent grade-band score scales are linked psychometrically for the ELA, mathematics and social studies, some tasks are used as linking tasks in each grade-band assessment that align with the extended standards in both adjacent grade-bands. All items in linking tasks are designed to be appropriate for students in both adjacent grade-bands.

The alignment studies (discussed in Chapter 8) confirm that all tasks in each grade-band, including linking tasks, align with ASMGs or extended standards for each separate grade-band and with the corresponding grade-band academic content standards.

Chapter 3: Spring 2011 Operational Test Administration

This section describes the spring 2011 operational test administration in the following areas:

- Student participation for the spring 2011 administration
- Demographics of participating students
- Test administration window, materials, and timelines
- Test administrator requirements
- Test administrator training
- Pre-assessment using the Student Placement Questionnaire
- Fidelity of administration and accuracy of scoring
- Test security provisions

Student Participation for the Spring 2011 Administration

Students participating in the spring 2011 operational administration were those students whose IEP team had determined that they met the following SC-Alt participation criteria for alternate assessment and who were ages 8–13 or 15 on September 1, 2010. These are the ages of typical students who are in grades 3–8 and 10.

- The student demonstrates a significant cognitive disability and adaptive skills, which result in performance that is substantially below grade-level achievement expectations even with the use of accommodations and modifications.
- The student accesses the state-approved curriculum standards at less-complex levels and with extensively modified instruction.
- The student has current adaptive skills requiring extensive direct instruction and practice in multiple settings to accomplish the application and transfer of skills necessary for application in school, work, home, and community environments.
- The student is unable to apply or use academic skills across natural settings when instructed solely or primarily through classroom instruction.
- The student's inability to achieve the state grade-level achievement expectations is not the result of excessive or extended absences or social, cultural, or economic differences.

Exhibit 3.1 indicates the age ranges of students who participated in the SC-Alt in spring 2011.

Exhibit 3.2 indicates the alternate assessment eligibility categories that were placed in each eligible student's state precoding file (precoding files enabled SCDE and AIR to ensure that the appropriate SC-Alt materials were delivered to teachers in time for the spring 2011 administration).

Exhibit 3.1: Age Reference Sheet for 2010–2011 Alternate Assessment, Spring 2011 Operational Administration

Age as of 9/1/10	Corresponding Birth Date Range		Test Required 2010–2011	Precode AA Eligibility Code
	Beginning DOB	Ending DOB		
5	09/02/04	09/01/05	none	5
6	09/02/03	09/01/04	none	5
7	09/02/02	09/01/03	none	5
8	09/02/01	09/01/02	SC-Alt Elem	2
9	09/02/00	09/01/01	SC-Alt Elem	2
10	09/02/99	09/01/00	SC-Alt Elem	2
11	09/02/98	09/01/99	SC-Alt Middle	3
12	09/02/97	09/01/98	SC-Alt Middle	3
13	09/02/96	09/01/97	SC-Alt Middle	3
14	09/02/95	09/01/96	none	5
15	09/02/94	09/01/95	SC-Alt HS	4
16	09/02/93	09/01/94	none	5
17	09/02/92	09/01/93	none	5
18	09/02/91	09/01/92	none	5
19	09/02/90	09/01/91	none	5
20	09/02/89	09/01/90	none	5
21	09/02/88	09/01/89	none	5

Exhibit 3.2: Precode Project Coding (Alternate Assessment Eligibility Field)

Code	SASI Dropdown List Description	Full Description
0	Criteria not met	The student does not meet criteria for alternate assessment.
2	SC-Alt Elem School	The student requires alternate assessment and meets the age eligibility requirement for assessment with the SC-Alt Elem School form this current school year (8–10 years old on September 1, 2010).
3	SC-Alt Middle School	The student requires alternate assessment and meets the age eligibility requirement for assessment with the SC-Alt Middle School form this current school year (11–13 years old on September 1, 2010).
4	SC-Alt High School	The student requires alternate assessment and meets the age eligibility requirement for assessment with the SC-Alt High School form this current school year (15 years old on September 1, 2010).
5	AltAssess NotAgeElig	The student requires alternate assessment but does not meet the age eligibility requirements to be assessed with SC-Alt this current school year (i.e., the student was younger than eight years, age 14, or older than 15 years on September 1, 2010).

Demographics of Participating Students

This section describes the demographics of participating students by test form (elementary, middle, or high school). Exhibit 3.3 presents the student demographics for participating students in each grade-band.

For the purpose of this report, the inclusion of students was based on the same criteria applied in the reporting of student scores. A student was included if the following criteria were met: (1) a signed security affidavit was received for the student, (2) the student was not noted to be excluded from reporting for some other reason (e.g., inappropriate administration procedures), and (3) the number of coded responses met the attemptedness requirement for student scoring (i.e., five valid responses) in at least one content area. The population of students reported, therefore, includes 1,486 elementary school test forms, 1,326 middle school test forms, and 355 high school test forms.

According to the attemptedness requirements, a student's responses to a test form could be assigned to one of four completion status categories: completion ("student satisfied attemptedness rule"), invalid due to too few scored responses ("student did not satisfy attemptedness rule"), invalid due to test administration errors ("test administrator did not follow instructions for starting tasks"), or not tested ("student did not answer any content area items"). For all content areas, the majority of students reported completed the administered test form; 99% or more of the eligible students completed ELA and mathematics, 69% to 70% completed science and social studies in the elementary and middle school grade-bands,² and 99% completed the high school biology assessment. Of the remaining student records, less than 1% of reported test forms were categorized as not tested or not meeting the attemptedness criteria.

Given that the number of students to be assessed on the high school test form was approximately one-third the number of students assessed on either the elementary or the middle school forms, the proportion of demographic characteristics of the student population was relatively consistent across grade-bands. In terms of ethnicity, African American students made up 49% to 52% of the assessed students across grade-bands; white students accounted for 41% to 44% of the students across grade-bands; and Hispanic students accounted for 1% to 6% of students across forms. Other ethnicities each accounted for less than 3% of the assessed population. Gender was also consistent across grade-bands with approximately a two-to-one ratio of male students (68%) to females (32%).

The classification of students in terms of English language proficiency was also consistent across grade-bands. The majority of students (96% to 99%) were classified as "English Speaker II," meaning that they had never been coded as an ESL student. The remaining language proficiency classifications each accounted for less than 1% of students by grade-band with the exception of "Pre-functional" (1% to 4%), indicating that the student scored pre-functional on the English language proficiency assessment and was receiving English as a second language (ESL) services. The percentage of pre-functional ESL students decreased across grade-bands.

The grade reported for a student in the school's database is the grade reported for funding purposes (EFA grade) and is often determined by the location of the student's educational program instead of by the student's age or years in school. Therefore, approximately 9% of students administered the elementary form (for students ages 8–10, the typical ages of students in grades 3–5) had reported EFA grades lower than grade 3 or higher than grade 5, with most of these students classified in the adjacent grades of 2 and 6. Of students administered the middle school form (for students ages 11–13, the typical ages for grades 6–8), 20% of the students were reported at grades below grade 6 or above grade 8. The vast majority of these students were

² Not all students were required to complete the science and social studies subject areas.

classified as grade 5 students (17% of all middle school form students), which indicates that these students were being served in educational programs housed in elementary schools. Of the students administered the high school form (for students age 15), 77% were reported as grade 9 or grade 10 (37% and 40%, respectively). Twelve percent (12%) of the high school form students were reported as grade 8 students, indicating that these students were being served in educational programs housed in middle schools. The purpose of assigning SC-Alt grade-band forms by age is to ensure that students are instructed and assessed on the appropriate grade-band curricula regardless of where their educational programs are housed.

The percentage of students receiving free lunch at schools decreases slightly across forms (67% to 63%), and the percentage of students receiving reduced-price meals is approximately the same across forms (6% to 7%). One student was indicated as being a migrant student; no students were indicated as being home-schooled. Thirteen elementary school students (less than 1%) were indicated as being medically homebound, as were 19 middle school students (1%) and eight high school students (2%).

Fourteen different disability codes were reported for students assessed with the SC-Alt. The coding system allowed students to be coded with more than one disability code. Students with the primary disabilities of severe mental disability, moderate mental disability, mild mental disability, and autism made up 82% to 91% of the students assessed with the SC-Alt. Of these, the percentage of students coded as having moderate mental disability increased across test forms (22% to 37%), while autism decreased from 25% in elementary school to 18% in high school. The rates of both severe mental disability and mild mental disability stayed about the same (9%–10% and 26%–27%, respectively). Although a few students were given a primary disability code of speech or language impairment, the vast majority of students received this code because they were receiving speech/language therapy as a supplementary service.

Exhibit 3.3: Summary of Demographic Information

	Grade-Band 3–5		Grade-Band 6–8		Grade 10	
	N	%	N	%	N	%
STUDENT'S ETHNICITY						
African American	722	48.6	683	51.5	179	50.4
American Indian	4	0.3	3	0.2	1	0.3
Asian	22	1.5	11	0.8	5	1.4
Hawaiian/Pacific Islander	2	0.1	1	0.1	1	0.3
Hispanic	89	6.0	50	3.8	5	1.4
Other	40	2.7	30	2.3	7	2.0
White	607	40.9	548	41.3	157	44.2
STUDENT'S GENDER						
Female	478	32.2	428	32.3	115	32.4
Male	1008	67.8	898	67.7	240	67.6
ESL (LANGUAGE)						
Advanced	.	0	.	0	.	0
Advanced Waiver	.	0	.	0	.	0

	Grade-Band 3–5		Grade-Band 6–8		Grade 10	
	N	%	N	%	N	%
Beginner	3	0.2	2	0.2	.	0
Beginner Waiver	.	0	1	0.1	.	0
English Speaker I	1	0.1	2	0.2	1	0.3
English Speaker II	1421	95.6	1284	96.8	350	98.6
Initially English Proficient	.	0	.	0	.	0
Intermediate	.	0	.	0	.	0
Intermediate Waiver	.	0	.	0	.	0
Pre-functional	60	4.0	35	2.6	4	1.1
Pre-functional Waiver	1	0.1	.	0	.	0
Title III First Year Exited	.	0	.	0	.	0
Title III Second+ Year Exited	.	0	.	0	.	0
Unknown	.	0	2	0.2	.	0
ELIGIBLE FOR FREE OR REDUCED-PRICE LUNCH						
Free Meals	993	66.8	844	63.7	223	62.8
Full-Pay Meals	390	26.2	384	29.0	111	31.3
Reduced Meals	103	6.9	98	7.4	21	5.9
Unknown	.	0	.	0	.	0
EFA GRADE (REPORTED GRADE FOR FUNDING)						
1	7	0.5	1	0.1	.	0
2	101	6.8	2	0.2	1	0.3
3	498	33.5	.	0	.	0
4	532	35.8	24	1.8	1	0.3
5	328	22.1	220	16.6	10	2.8
6	14	0.9	443	33.4	3	0.9
7	.	0	382	28.8	5	1.4
8	3	0.2	235	17.7	43	12.1
9	3	0.2	16	1.2	131	36.9
10	.	0	2	0.2	142	40.0
11	.	0	1	0.1	15	4.2
12	.	0	.	0	4	1.1
COMPLETION STATUS: Attempted						
ELA	1485	99.9	1326	100	355	100
Math	1485	99.9	1320	99.6	350	98.6
Science/Biology	1031	69.4	914	68.9	350	98.6
Social Studies	1024	68.9	922	69.5	.	0
COMPLETION STATUS: Not Tested						
ELA	1	0.1	.	0	.	0
Math	1	0.1	2	0.2	4	1.1

	Grade-Band 3–5		Grade-Band 6–8		Grade 10	
	N	%	N	%	N	%
Science/Biology*	455	30.6	412	31.1	5	1.4
Social Studies*	460	31.0	403	30.4	355	100
Completion Status: Test Not Valid—Student Received Fewer Than Five Scored Responses						
ELA	.	0	.	0	.	0
Math	.	0	4	0.3	.	0
Science/Biology	.	0	.	0	.	0
Social Studies	2	0.1	1	0.1	.	0
Completion Status: Test Not Valid—Test Administrator Did Not Follow Instructions for Starting Tasks						
ELA	.	0	.	0	.	0
Math	.	0	.	0	1	0.3
Science/Biology	.	0	.	0	.	0
Social Studies	.	0	.	0	.	0
Special School Status Fields						
Migrant Status	1	0.1	.	0	.	0
Home-Schooled	.	0	.	0	.	0
Medical Homebound	13	0.9	19	1.4	8	2.3
IEP Disability Codes (Multiple Codes per Student)						
Severely Mentally Disabled	136	9.2	134	10.1	31	8.7
Moderately Mentally Disabled	332	22.3	404	30.5	132	37.2
Mildly Mentally Disabled	387	26.0	364	27.5	97	27.3
Autism	364	24.5	270	20.4	63	17.8
Deaf/Blindness	.	0	.	0	.	0
Emotional Disability	15	1.0	10	0.8	1	0.3
Hearing Impaired	21	1.4	21	1.6	11	3.1
Learning Disability	44	3.0	27	2.0	3	0.9
Multiple-Disability	1	0.1	2	0.2	2	0.6
Other Health Impaired	78	5.3	63	4.8	14	3.9
Orthopedically Impaired	61	4.1	48	3.6	14	3.9
Speech or Language Impaired	1085	73.1	661	49.9	118	33.2
Traumatic Brain Injury	11	0.7	8	0.6	4	1.1
Visually Impaired	53	3.6	48	3.6	11	3.1
TOTAL	1486	100	1326	100	355	100

*Not all students were required to complete the science and social studies subject areas.

Test Administration Window, Materials, and Timelines

The spring 2011 administration of the SC-Alt included the following important dates:

- SC-Alt test administration training for teachers new to the SC-Alt operational administration (did not administer in 2009 or 2010), five regional SCDE workshops: January 10–14, 2011
- District-level SC-Alt test administration training for all test administrators: January 31–February 28, 2011
- Test materials arrived in district: February 24, 2011
- Assessment window: March 7–April 29, 2011
- Teachers returned materials to the district test coordinator for alternate assessment (DTC-Alt): May 4, 2011
- Materials received by contractor: May 6, 2011

Teachers had approximately eight weeks to review the materials and complete the test administration. Teachers received both printed and physical manipulatives to use during test administration. They were also responsible for collecting a few common classroom items that were familiar to the student to use with several tasks.

Test Administrator Requirements

Test administrators were required to receive training on all phases of the administration of the SC-Alt and had to be one of the following:

- A certified employee of the district
- An employee of the district who is a critical needs teacher and has a letter of eligibility, an interim certificate, or a critical needs certificate
- A substitute teacher who is certified and employed by the district on an as-needed basis
- Someone who was a certified teacher but has allowed the teaching certificate to expire owing to retirement, change of career, or some other reason and has been approved by the district test coordinator or the DTC-Alt as a qualified test administrator
- Someone who is not certified but has been employed by the school district in an instructional capacity and has been approved by the DTC-Alt as a qualified test administrator

If a test was administered in a location other than the school, the test administrator still had to meet the criteria specified above.

Test Administrator Training

Test administration training was required for all test administrators. The SC-Alt is individually administered with a standard script and scored by the test administrator as the assessment is

being conducted. Fidelity of administration and scoring is essential to the validity of the assessment results.

Teachers who administered the SC-Alt during spring 2011 but who did not administer the SC-Alt in spring 2009 or 2010 were required to attend a SCDE training session. In addition, all teachers who administered the SC-Alt in spring 2011, including those who attended the SCDE workshops, were required to attend a district-level SC-Alt administration training session conducted by the DTC-Alt. At the completion of the training sessions, each test administrator was required to sign and submit to SCDE an acknowledgment of receiving training and readiness to conduct the assessment.

The training included the following elements:

- Review of the eligibility criteria for students participating in the alternate assessment
- Overview of the Extended Standards, emphasizing the link to the general education standards
- Explanation of how the assessment was developed, including the role of the review committees
- Review of test administrator requirements, test security, and test materials
- Training and practice in pre-assessment using the SPQ
- Description of the assessment format and procedures:
 - Setup
 - Script
 - Scoring
 - Adaptive instructions
- Instruction for making SC-Alt tasks accessible
- Overview of assistive technology and the alternate assessment
- Administration and scoring instruction and practice using released test items provided on video clips of South Carolina teachers administering a task to students representing a variety of disabilities and ethnicities
- Scoring qualifying round
- Review of procedures for receiving and shipping materials back to the DTC-Alt

Pre-Assessment Using the Student Placement Questionnaire

As noted earlier in this Technical Report, the SC-Alt uses the SPQ as a pre-assessment instrument to determine the most appropriate starting point in the assessment. Recall that the SPQ requires the teacher to evaluate the student on 12 to 15 “can do” statements addressing the student’s skills and knowledge in each content area on the basis of the teacher’s prior instructional knowledge of the student. A total score computed from the teacher’s SPQ responses indicates the initial starting task for the assessment. Once the assessment has begun, the test

administrator is required to adjust the starting point for the student if the student is not successful on the first task. Rules have been established for adjusting the starting tasks and for determining when the assessment should be concluded. The starting and stopping rules used with the SPQs for the 2011 administration are presented in Appendix B.

Fidelity of Administration and Accuracy of Scoring

During the assessment administration, a monitor had to be present to observe all assessment sessions and verify the use of proper assessment procedures and the authenticity of student responses. Monitors had to be trained, and they had to sign a Test Administrator Security Affidavit to verify that the appropriate procedures were used. The Test Administrator Security Affidavit is located in the answer folder and includes the principal's verification of the use of appropriate assessment and scoring procedures. Whenever the requested signatures were missing, the administration was considered an invalid administration.

Some of the assessments (6% to 15%) were audited by having trained raters score the student's performance independently, either from a videotaped recording or while witnessing the assessment directly. The results of these studies are reported in detail in Appendix C.

Test Security Provisions

This section describes the test security procedures associated with the SC-Alt. SCDE has the following test security measures in place:

- Each local school board must develop and adopt a district test security policy. The policy must provide for the security of the materials during testing and the storage of all secure tests and test materials before, during, and after testing. Before and after testing, all materials must be stored at a location(s) in the district under lock and key.
- Each District Superintendent must designate annually one individual in each district for each mandated assessment who will be the sole individual in the district authorized to procure test instruments that are used in testing programs administered by or through the State Board of Education. The designated individual for alternate assessment is the DTC-Alt. The DTC-Alt is responsible for receiving and distributing all SC-Alt materials and ensuring that all SC-Alt administration procedures and requirements are met.
- All school and district personnel who may have access to SC-Alt test materials or to the location in which the materials are securely stored must sign the Agreement to Maintain Test Security and Confidentiality before they are given access to the materials.
- Test administrators must be trained annually to administer the SC-Alt and must meet all test administrator requirements.
- An assessment monitor must observe all assessment sessions and verify the use of proper assessment procedures and the authenticity of student responses for each completed assessment.
- Test administrators must complete an SC-Alt Test Administrator Security Affidavit for each student they assess.

Chapter 4: Setting Performance Standards

For the South Carolina Alternate Assessment, two standard-setting workshops were conducted: In June 2007, the first standard-setting workshop convened a diverse panel of 105 educators, parents, and educational administrators to recommend status performance standards based on the spring 2007 operational test administration data for ELA, mathematics, and science, and for the field-test data for social studies. In the second workshop, in September 2010, 19 panelists recommended standards for high school biology based on spring 2010 field-test data. This chapter summarizes the descriptions of achievement levels, the procedures used for setting standards for each content area, and the recommended standards themselves, including student impact information. This is followed by an impact analysis of the biology standards, based on the spring 2011 operational administration data. Complete details of the two standard-setting workshops can be found in separate reports (American Institutes for Research, 2007; American Institutes for Research and South Carolina Department of Education, 2010b).

Using the Item Descriptor (ID) Matching method (see Cizek & Bunch, 2007; Ferrara, Perie, & Johnson, 2008), the panelists reviewed test items and the corresponding Descriptions of Achievement Levels (DALs) and then recommended performance standards for Level 2, Level 3, and Level 4 achievement levels. These standards were translated into cut points on the student proficiency scale by AIR psychometricians. This section describes the process and outcomes of the standard-setting workshop.

Descriptions of Achievement Levels

DALs are key elements in standard-setting processes. DALs define the content area knowledge, skills, and processes that examinees at a performance level are expected to possess. The descriptions of Level 1, Level 2, Level 3, and Level 4 performance that SCDE developed make up the public statement about what and how much South Carolina educators want students to know and be able to do for each grade level and content area. Level 3 and higher represents “proficient performance” for NCLB reporting.

The development of the DALs for ELA, mathematics, science, and social studies followed a multistep process involving AIR staff and SCDE staff working with committees of teachers, parents, and special education administrators. The process was begun by examining the DALs used with the other South Carolina assessment programs (PACT, HSAP, PACT-Alt, and HSAP-Alt) and the performance-level descriptors for alternate assessments used by other states. During spring 2007, these DALs were developed and refined over multiple meetings between AIR, SCDE, and stakeholder committees to determine what proficiency meant for students participating in each grade-band of the SC-Alt. Some additional refinement occurred during the standard-setting workshop in June 2007; the final version of these DALs was presented to the State Board of Education on September 12, 2007, and posted on the SCDE website.

The DALs for high school biology were written by AIR and reviewed by SCDE prior to the standard-setting meeting. A subcommittee of standard-setting panelists reviewed the biology DALs on the first day of the September 2010 standard-setting workshop.

In the SC-Alt standard-setting workshops, panelists used the DALs presented in Appendix D when they placed their cut scores.

The ID Matching Standard-Setting Process

The ID Matching standard-setting process, described in the standard-setting plans submitted to SCDE and reviewed by the South Carolina Technical Advisory Committee, was used at both standard-setting workshops in Columbia, SC (in June 2007 and September 2010). When standards were to be set in multiple subjects, the panels were divided into subject-specific groups. For subjects that were assessed in multiple grade-bands, anchor standards were first established in the lowest and highest grade-bands (e.g., grade-bands 3–5 and 10). AIR staff provided training and led the panelists through two rounds of ID Matching to set the Level 3 standard first, followed by the Level 2 and 4 standards.

Before the participants made each of their recommendations using the ID Matching procedure, they were given a readiness form to ensure that they fully understood the task and were prepared to place the performance standard. The participants indicated unanimously that they understood the task and were prepared to make performance standard recommendations.

Goals of the Standard Setting

The goals of the meeting, as stated to the panelists, were as follows:

- Recommend performance standards on the ELA, mathematics, science, social studies, and/or biology assessments that correspond to the DALs for Level 2, Level 3, and Level 4 performance levels
- Consider the agreement and impact data to guide judgments about item difficulty and placement of the performance standards
- Recommend to SCDE the appropriate placement of cut points on the student proficiency scales for each grade-band assessment

Panel Composition

Across the two workshops, 124 panelists participated in recommending performance standards across five content areas: ELA, mathematics, science, social studies, and biology. The overall composition of the panel followed the SCDE-provided specifications and was broadly designed to ensure that the panel was widely diverse and represented a cross-section of South Carolina's educators and non-educators.

Standard-Setting Workshop Activities

Workshop participants recommended performance standards for the assessments during two rounds of deliberation for each DAL in each content area and in each grade-band as follows.

- Set standards in anchor grade-bands (3–5 and 10)
 - Participants complete Rounds 1 and 2 for each performance-level standard.
 - Table leaders articulate standards across grades and content areas (align them on the basis of content considerations).
 - For the biology standard setting, table leaders moderate the standards with respect to existing performance standards in science.

- Set standards in intermediate grade-band (6–8), if needed.
 - Participants complete Rounds 1 and 2 for each performance-level standard.
 - Table leaders *articulate* standards across grades and content areas (align them on the basis of content considerations and consistency with anchor grade standards).

Throughout the workshops, the panelists had many opportunities to reflect on the pattern of performance standards they were recommending. Their general conclusion was that they were satisfied that the standards made sense from a content and experiential point of view. They felt that the patterns reflected the requirements of the content standards and the realities of student performance.

With few exceptions, panelists recommended standards that followed an orderly progression of increasing achievement across levels and grade-bands. Specifically, with the exception of mathematics at the grade-band 6–8 and grade 10, all recommended achievement-level standards increased in difficulty in subsequent grade-bands. Exhibits 4.1–4.5 show the scale score associated with the cut score recommended by each panel. These results were achieved through the process of setting cut scores at anchor grades, making sure that they resulted in consistent expectations across grade-bands, and providing articulated standards as a starting point for intermediate grade-bands.

Cut Score Review and the Setting of Final Cut Scores

The results of the standard-setting workshop for ELA, mathematics, science and social studies were presented to the Technical Advisory Committee (TAC) of the Office of Assessment, SCDE, on July 27, 2007. The TAC discussed the results of the standard-setting workshop, reviewed the articulation of the cut scores by grade level, and recommended strategies to the Office of Assessment staff for improving the articulation of the final scores while respecting and maintaining the basic cut score decisions made by the workshop panelists.

A committee of Office of Assessment staff examined the scale score articulation and the percentage of students in performance levels by grade and recommended minor adjustments to the original cuts made by the workshop panelists. The adjustments made to each cut score and the resulting final cut scores are presented in Exhibits 4.1–4.5. These cut-scores were approved by the State Superintendent of Education and were presented to the South Carolina State Board of Education on September 12, 2007.

In Exhibits 4.1–4.4, the combined standard error of the panelist-recommended cut score (labeled “SE 2007”) expresses the joint uncertainty of the IRT-based estimate of the conditional standard error of measurement at the cut score, together with the sampling error of the median agreement per cut score among panelists. The standard error of the median cut score agreement among the panelists, as suggested by Huynh (2003), is listed in column 6. However, two additional details about the standard errors of the median are important to note: First, the standard errors were based on the actual recommended cut scores, and any post hoc adjustment to the cut scores was treated as a constant adjustment. In other words, the adjusted cut score still had the same standard error. Second, the standard errors were initially calculated as standard errors of the page numbers in the ordered-item booklet and then transformed to the scale score metric.

The estimate of the conditional standard error of measurement depends on the set of items used at the time and on the distribution of operational item response patterns observed in a given administration. The entries of the right-most column of Exhibits 4.1–4.5 (labeled “CSEM 2011”) display the empirically estimated conditional standard error of measurement at the final, adjusted cut score. This latter estimate is taken from the spring 2011 operational data, computed as the root mean square standard error of the scale-score estimates within ± 5 scale units of the cut point. The CSEM 2011 values indicate the precision of the current test instrument at the final cut points determined earlier, in 2007 and 2010, respectively.

Exhibit 4.1: Panel Recommended and Adjusted Final Cut Scores—ELA

Performance Level	Panel Recommended Cut Scores		Adjustment to Final Cut Scores			2011 Conditional Standard Error of Measurement (CSEM 2011)
	Scale Score	2007 Combined Standard Error (SE 2007)	Level of Adjustment (± SE)	Final Cut Scale Score	Standard Error of Cut Scale Score	
Grade-Band 3–5						
Level 2	403	13.75	None	403	2.96	14.32
Level 3	466	9.54	None	466	1.59	9.69
Level 4	491	12.26	None	491	1.73	10.81
Grade-Band 6–8						
Level 2	417	9.64	None	417	3.81	13.01
Level 3	473	7.99	0.5	477	1.09	10.16
Level 4	501	9.18	None	501	1.45	11.13
Grade 10						
Level 2	429	10.56	None	429	3.38	12.18
Level 3	478	9.11	1	487	0.66	9.94
Level 4	503	9.68	1	514	1.77	10.59

Exhibit 4.2: Panel Recommended and Adjusted Final Cut Scores—Mathematics

Performance Level	Panel Recommended Cut Scores		Adjustment to Final Cut Scores			2011 Conditional Standard Error of Measurement (CSEM 2011)
	Scale Score	2007 Combined Standard Error (SE 2007)	Level of Adjustment (± SE)	Final Cut Scale Score	Standard Error of Cut Scale Score	
Grade-Band 3–5						
Level 2	423	10.22	-1	413	0.66	13.52
Level 3	476	9.59	None	476	0.21	10.93
Level 4	526	14.48	None	526	4.63	13.23
Grade-Band 6–8						
Level 2	425	10.18	None	425	0.50	13.55
Level 3	476	9	1.5	489	0.16	10.49
Level 4	529	10.46	0.5	534	0.74	11.69
Grade 10						
Level 2	434	11.93	None	434	2.19	14.66
Level 3	476	14.76	1.5	498	1.97	10.58
Level 4	528	13.19	1	541	3.82	11.80

Exhibit 4.3: Panel Recommended and Adjusted Final Cut Scores—Science

Performance Level	Panel Recommended Cut Scores		Adjustment to Final Cut Scores			2011 Conditional Standard Error of Measurement (CSEM 2011)
	Scale Score	2007 Combined Standard Error (SE 2007)	Level of Adjustment (± SE)	Final Cut Scale Score	Standard Error of Cut Scale Score	
Grade-Band 3–5						
Level 2	430	10.83	None	430	1.51	14.77
Level 3	474	10.36	-0.5	469	3.25	14.28
Level 4	496	10.38	None	496	0.81	13.83
Grade-Band 6–8						
Level 2	447	9.66	None	447	0.06	13.35
Level 3	484	9.61	0.5	489	0.50	12.23
Level 4	514	11.33	None	514	0.95	13.52

Exhibit 4.4: Panel Recommended and Adjusted Final Cut Scores—Social Studies

Performance Level	Panel Recommended Cut Scores		Adjustment to Final Cut Scores			2011 Conditional Standard Error of Measurement (CSEM 2011)
	Scale Score	2007 Combined Standard Error (SE 2007)	Level of Adjustment (± SE)	Final Cut Scale Score	Standard Error of Cut Scale Score	
Grade-Band 3–5						
Level 2	423	16.64	None	423	2.98	15.39
Level 3	485	14.39	0.5	492	11.93	13.23
Level 4	549	14	None	549	2.04	16.15
Grade-Band 6–8						
Level 2	439	14.04	None	439	5.96	14.51
Level 3	490	12.58	1.5	503	1.28	14.05
Level 4	560	26.91	None	560	10.57	18.11

The cut scores and standard errors for the high school biology assessment are given in Exhibit 4.5. This exhibit is structured the same as Exhibits 4.1–4.4, except that the combined standard error estimates are computed from 2011 operational data. The biology cut scores were approved by the State Superintendent of Education on December 20, 2010.

Exhibit 4.5: Final Cut Scores—High School Biology

Performance Level	Panel Recommended (High) Cut Scores		Adjustment to Final Cut Scores			2011 Conditional Standard Error (2011 CSEM)
	Scale Score	2011 Combined Standard Error (SE 2011)	Level of Adjustment (± SE)	Final Cut Scale Score	Standard Error of Cut Score	
Grade 10						
Level 2	408	27.3	None	408	5.61	26.48
Level 3	484	35.4	None	484	8.83	24.18
Level 4	519	28.1	None	519	4.91	26.51

For ELA, mathematics, science, social studies, and biology, the final cut scores, the percentage of students performing at each performance level, and the cumulative percentage of students at or above each level in the spring 2011 operational administration are presented in Exhibits 4.6–4.10.

Exhibit 4.6: Percentage of Students at Each Performance Level—ELA

Performance Level	Scale Score Cut Score	Percentage in Level (2007 Data)	Cumulative Percentage (at and above) for Each Performance Standard (2007 Data)
Grades 3–5			
Level 1	—	12.6 %	100.0 %
Level 2	403	25.4 %	87.4 %
Level 3	466	21.9 %	62.0 %
Level 4	491	40.1 %	40.1%
Grade-Band 6–8			
Level 1	—	12.9 %	100.0 %
Level 2	417	23.3 %	87.2 %
Level 3	477	14.9 %	63.9 %
Level 4	501	49.0 %	49.0 %
Grade 10			
Level 1	—	13.4 %	100.0 %
Level 2	429	23.6 %	86.6 %
Level 3	487	12.5 %	63.1%
Level 4	514	50.6 %	50.6 %

Exhibit 4.7: Percentage of Students at Each Performance Level—Mathematics

Performance Level	Scale Score Cut Score	Percentage in Level (2007 Data)	Cumulative Percentage (at and above) for Each Performance Standard (2007 Data)
Grade-Band 3–5			
Level 1	—	14.3 %	100.0 %
Level 2	413	30.8 %	85.7 %
Level 3	476	29.3 %	54.9 %
Level 4	526	25.7 %	25.7 %
Grade-Band 6–8			
Level 1	—	15.9 %	100.0 %
Level 2	425	28.5 %	84.1%
Level 3	489	25.9 %	55.6 %
Level 4	534	29.8 %	29.8 %
Grade 10			
Level 1	—	16.1 %	100.0 %
Level 2	434	30.1 %	84.0 %
Level 3	498	28.9 %	53.9 %
Level 4	541	24.9 %	24.9 %

Exhibit 4.8: Percentage of Students at Each Performance Level—Science

Performance Level	Scale Score Cut Score	Percentage in Level (2007 Data)	Cumulative Percentage (at and above) for Each Performance Standard (2007 Data)
Grade-Band 3–5			
Level 1	—	19.8 %	100.0 %
Level 2	430	18.2 %	80.2 %
Level 3	469	17.5 %	62.0 %
Level 4	496	44.5 %	44.5 %
Grade-Band 6–8			
Level 1	—	22.1 %	100.0 %
Level 2	447	18.5 %	77.9 %
Level 3	489	15.3 %	59.3 %
Level 4	514	44.0 %	44.0 %

Exhibit 4.9: Percentage of Students at Each Performance Level—Social Studies

Performance Level	Scale Score Cut Score	Percentage in Level (2007 Data)	Cumulative Percentage (at and above) for Each Performance Standard (2007 Data)
Grade-Band 3–5			
Level 1	—	19.3 %	100.0 %
Level 2	423	32.7 %	80.7 %
Level 3	492	30.1 %	48.1 %
Level 4	549	18.0 %	18.0 %
Grade-Band 6–8			
Level 1	—	19.7 %	100.0 %
Level 2	439	27.3 %	80.3 %
Level 3	503	34.1 %	53.0 %
Level 4	560	19.0 %	19.0 %

Exhibit 4.10: Percentage of Students at Each Performance Level—Biology

Performance Level	Scale Score Cut Score	Percentage in Level (2011 Data)	Cumulative Percentage (at and above) for Each Performance Standard (2011 Data)
Grade 10			
Level 1	—	10.3%	100.0%
Level 2	408	27.4%	89.7%
Level 3	484	19.1%	62.3%
Level 4	519	43.1%	43.1%

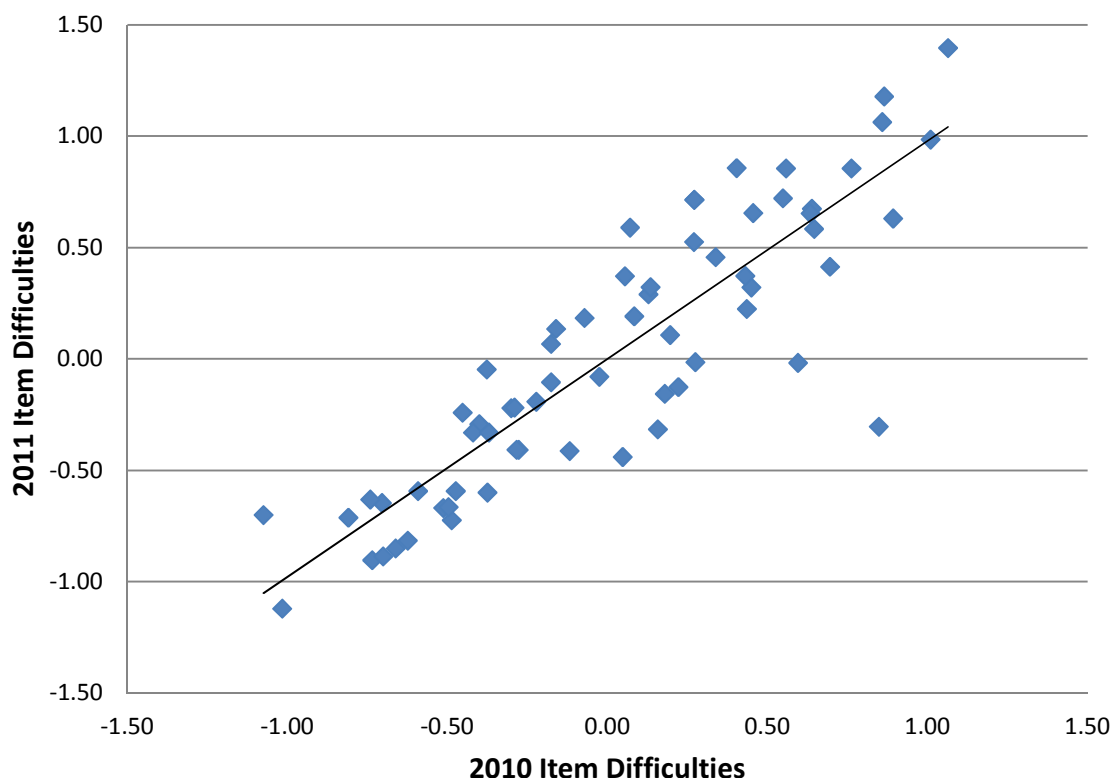
Consistency of Item Parameter Estimates between the 2010 Biology Field Test and 2011 Operational Assessments

Sampling error, changes in the student population and curricular changes can cause item parameters to differ between test administrations, which could potentially threaten the reliability of the assessment. It is therefore good practice to monitor the item difficulties of an assessment, particularly when the instrument is employed operationally for the first time.

The item difficulty scatter plot in Exhibit 4.11 displays the scaled item difficulty parameters estimated from the 2011 operational data against the item difficulty estimates from the 2010 field test. Most points fall along the 45-degree diagonal line, indicating appreciably high stability of item parameter estimates. There is one outlier, an item (item ID 1459) that became considerably easier for the 2011 operational administration. The Pearson correlation between the 2010 and 2011 item parameter estimates is 0.88 with all items included, and 0.91 after removing the outlier. The Spearman rank correlation with the item included is 0.91. These statistical results are in support of general stability of the biology item parameters.

The outlying item (ID 1459) addresses the genetic inheritance of biological traits of animals in contrast to shared environmental conditions. At the time of this writing, it remains unclear what caused the change in item difficulty. The possibility that biology instruction has improved in the 2010–2011 school year would be a promising candidate for investigation.

Exhibit 4.11: Scaled 2011 Operational Item Difficulty Estimates Plotted Against 2010 Field-Test Estimates



Number of Tasks Administered in the Biology Field Test and Operational Administration

The SC-Alt is a form-adaptive, not a fixed-form test. In the assessment, students are administered subsections of the form, according to how generally competent in biology they were rated by their teachers (or how competent in science for the 2010 biology field test). The 2010 field-test assessment would have a minimum of nine or ten tasks, depending on start point, and a maximum of 12.³ The 2011 operational assessment had a minimum of six or seven tasks, depending on start point, and a possible maximum of 12.⁴ How many tasks were actually administered to the students is detailed in Exhibit 4.12.⁵ In the 2010 field test, 87% of the students were administered the minimum of nine or more tasks. In the 2011 operational assessment, 99% of the students were administered the minimum of six or more tasks.

Exhibit 4.12: Number of Tasks Administered

Number of Task	2010 FT		2011 OT	
	N	%	N	%
1	0	0.0	1	0.3
2	0	0.0	0	0.0
3	1	0.2	0	0.0
4	1	0.2	1	0.3
5	19	4.1	1	0.3
6	5	1.1	70	20.0
7	34	7.3	118	33.7
8	3	0.6	22	6.3
9	96	20.6	13	3.7
10	198	42.5	54	15.4
11	5	1.1	6	1.7
12	104	22.3	64	18.3

Scale Score Distributions of the Biology Field Test and Operational Administration

To facilitate the comparison between the 2010 field test and the 2011 operational assessment results, Exhibit 4.13 overlays the scale score distributions of the two administrations. Both

³ 2010 FT: There are two possible start points, task 1 and task 3. When the assessment starts at task 1, then at least the first nine tasks are to be administered. If the student is still successfully performing at task 9, then the assessment should be continued up to a possible maximum of 12 tasks. Otherwise, if the assessment starts at task 3, then all 10 tasks (tasks 3–12) are to be administered.

⁴ 2011 OT: There are three possible start points, task 1, task 3, and task 6. When the assessment starts at task 1, then at least the first six tasks need to be administered. If the student is still successfully performing at task 6, then the assessment should be continued, up to a possible maximum of 12 tasks. When the assessment starts at task 3, then all seven tasks from 3 to 9 must be administered. If the student still successfully performs at task 9, then the assessment should be continued up to a possible maximum of 10 tasks (tasks 3–12). Otherwise, the assessment should start at task 6, and all seven tasks (6–12) should be administered.

⁵ This table counts the number of tasks administered independent of the starting task. In particular, a few cases of non-standard starting tasks are included in this tabulation.

distributions are centered at approximately the same mean, but the 2011 operational administration has more students with scale scores at the low and high ends of the reportable scale score range. This results in a larger variance of the 2011 operational scale score distribution than on the 2010 field test

Exhibit 4.13: Scale Score Distributions of the 2010 FT and 2011 Operational Biology Assessment

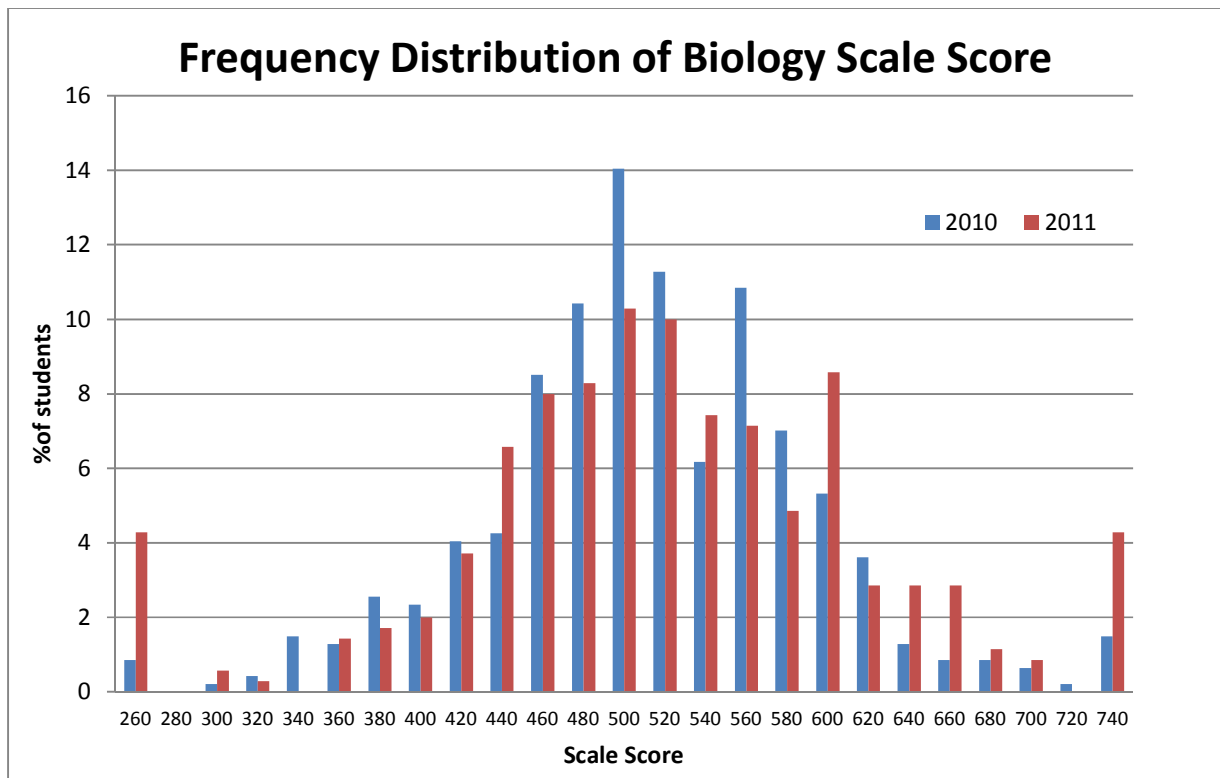


Exhibit 4.14 shows the univariate summary statistics of the scale score distributions of the 2011 operational biology assessment and the 2010 field test, respectively. As already indicated by the histogram of Exhibit 4.13, the data of the two administrations have similar means, but the distribution of the 2011 operational data shows a somewhat larger dispersion, due to more students scoring at the two extreme limits of the scale. The disaggregated means of male and female and black and white students are remarkably similar for the 2011 operational administration data.

Exhibit 4.14: Distribution Statistics, 2010 FT and 2011 OT Biology Assessments

	2010 FT			2011 OT		
	N	Mean	SD	N	Mean	SD
Total	466	502.9	79.5	350	509.4	100.5
Female	157	510.4	80.8	112	509.8	100.8
Male	309	499.0	78.7	238	509.3	100.6
Black	237	503.9	83.9	178	509.5	102.7
White	207	502.4	75.4	153	508.9	99.2

Biology Impact Data, Overall and by Demographic Subgroups

Exhibit 4.15 presents comparative data for the cumulative percentages of students in the performance levels for the 2011 operational vs. the 2010 biology field test administrations.⁶ The total percentage of students in performance Levels 3 and 4 was 62.3% for the 2011 operational administration, quite similar to the 61.6% obtained with the 2010 field-test data.

Exhibit 4.15: SC-Alt Biology 2011 vs. Field-Test Results

2011 High School Biology OT					
Percent in Performance Level or Higher	Total	Female	Male	Black	White
1	100.0	100.0	100.0	100.0	100.0
2	89.7	90.2	89.5	87.6	91.5
3	62.3	63.4	61.8	64.6	58.8
4	43.1	46.4	41.6	46.6	38.6
N	350	112	238	178	153
2010 High School Biology FT					
Percent in Performance Level or Higher	Total	Female	Male	Black	White
1	100.0	100.0	100.0	100.0	100.0
2	90.1	89.8	90.3	90.3	90.3
3	61.6	63.7	60.5	61.2	62.8
4	40.0	47.1	36.3	41.4	39.6
N	466	157	309	237	208

⁶ The figures in Exhibits 4.15 and 4.16 reflect the exact empirical impact data from the two administrations. There is a slight discrepancy between these and the figures reported in the standard-setting technical report. For instance, Exhibit 4.15 shows 40% of the field-tested students in performance level 4, while the standard-setting technical report listed this figure as 38.7%. The latter figure presents projection data—expected impacts modeled from item parameters and an assumed Normal-shaped latent ability distribution.

Impact Data Comparison to 2010 SC-Alt Physical Science Scores

2011 Biology impact data are compared with 2010 Physical Science data in Exhibit 4.16. Generally, more students reached performance Levels 3 and 4 in Biology (62.3%) than in Physical Science (54.2%). Both high school biology and physical science data show very similar impact percentages across gender and ethnic subgroups.

Exhibit 4.16: 2011 HS Biology vs. 2010 HS Physical Science Impact Data

2011 High School Biology OT					
Percent in Performance Level or Higher	Total	Female	Male	Black	White
1	100.0	100.0	100.0	100.0	100.0
2	89.7	90.2	89.5	87.6	91.5
3	62.3	63.4	61.8	64.6	58.8
4	43.1	46.4	41.6	46.6	38.6
N	350	112	238	178	153
2010 High School Physical Science					
Percent in Performance Level or Higher	Total	Female	Male	Black	White
1	100.0	100.0	100.0	100.0	100.0
2	76.4	80.3	74.4	73.7	78.5
3	54.2	59.0	51.7	56.8	52.4
4	32.6	35.3	31.2	36.8	28.9
N	356	122	234	190	149

Chapter 5: Technical Characteristics and Interpretation of Student Scores

This section describes the psychometric analyses conducted as part of the South Carolina Alternate Assessment (SC-Alt) 2007–2011 operational administrations. In 2011, embedded field-test tasks and items in ELA, mathematics, science, and social studies were newly calibrated and evaluated. In addition, the science forms also contained two operational field-test tasks. In spring 2011, no field-testing was performed in biology.

In order to provide a complete description of the technical characteristic of the 2011 assessment in all content areas, this chapter also reports the data analysis results for the sections of the assessment that had previously been calibrated using the 2007, 2008, and 2009 operational data and the independent field test for high school biology (see American Institutes for Research and South Carolina Department of Education, 2008, 2009, 2010a). The reported analyses are intended to ensure the quality of the items, the assessment materials and instruments, and the score reporting scales as measures of state academic standards.

As a reminder to the reader, there are three grade-band forms in each content area: elementary school (grades 3–5), middle school (grades 6–8), and high school (grade 10). ELA and mathematics are assessed on each grade-band, physical science and social studies only at elementary and middle school grades, and biology is assessed only at high school level. At each grade-band, the assessments have three potential starting tasks that correspond to three levels of task complexity (low, moderate, and high). Students are assigned to a starting task on the basis of teacher judgments recorded in the Student Placement Questionnaire (SPQ) for each content area. Linking tasks connect the grade-band forms so that the vertical test scale could be created.

Analysis and Scaling of Items, Tasks, and Test Forms

The ELA, mathematics, science, social studies, and biology assessments underwent comprehensive psychometric analyses, including initial item calibrations, after their earlier field-testing. Final calibrations were estimated for the ELA, mathematics, and science content areas on the basis of operational data gathered during the spring 2007 operational administration; final calibrations for social studies were computed from operational data from the spring 2008 administration. Calibrations based on operational data were considered superior to those based on field-test data. The vertical scales were also defined using the linking tasks as the vehicle that connected the elementary, middle, and high school forms. High school biology was field-tested in 2010 and administered operationally in spring 2011.

AIR calibrated the operational items, estimated examinee proficiencies, and calculated scale scores and achievement levels for operational forms. This process entailed examining item statistics to ensure quality measurement across the range of the assessment, calibrating the items within each content area to a common scale, and then applying a maximum-likelihood (ML) scoring algorithm to each student's responses to estimate his or her proficiency scores and assign the correct achievement level.

Assignment of Examinees to Starting Tasks and Item Calibration and Test Forms Linking

All eligible students participated in the spring 2011 test administrations. The sample sizes of approximately 1,486 students in elementary, 1,326 in middle school, and 355 in high school, per content area, enabled effective calibration across task starting points and grade-bands. Students

were assigned to one of three starting points on the basis of the sum of the teacher responses on the SPQ. The SPQ cut scores were shown to correlate with student achievement scores on the 2006 field-test administrations (for details, see American Institutes for Research, 2008). The assignment of student starting tasks based on the SPQ cut scores was intended to expose students to items that were ideally suited to their current level of achievement while ensuring that (a) each student responded to an adequate number of items so that reliable and content-valid proficiency scores could be estimated and (b) an adequate number of students responded to each item for the joint calibration to be reliable.

Teachers were instructed to administer all tasks associated with the assigned starting point, with provisions for dropping to a lower starting point (task) if the student was unable to respond to the items in the task at the assigned starting point. Students who were assigned to high and moderate levels of the assessment but were unable to respond to items in the tasks at those levels may have been moved back to a less difficult starting point.

The linking design allowed a joint (concurrent) calibration of all items within a content area and the placement of the items on a common difficulty scale. The tasks actually used to link the grade-band forms (linking tasks) were selected, in part, on the basis of their moderate difficulty levels. Moderately difficult tasks contribute to more stable linking across levels than tasks that may be either too easy or too difficult for the examinees.

Linking across grade-band forms was accomplished by using common tasks across grade-bands. Some of the tasks from the elementary form are on the middle school form; some of the tasks from the middle school form are on the high school form. For the 2011 assessments, linking tasks were only used for the ELA, mathematics, and social studies content areas. In general, tasks are assigned in such a manner that the forms increase slightly in difficulty as examinees progress through the grade-bands. This means that a task assigned to the moderate level of complexity in the elementary form may be assigned to the low-moderate level in the middle school form.

See Appendix E for a summary of the linking design in each of the four content areas.

Analysis Plan

AIR's analyses presented in the remainder of this chapter were conducted in five steps:

1. Data preparation and quality control
2. Classical item analysis
3. Review of items not meeting psychometric criteria for inclusion on operational forms
4. Joint calibration of items according to the Rasch model
5. Final achievement estimation and scale score calculation for operational forms

Data Preparation and Quality Control

Before analyzing the operational test data, AIR psychometricians performed a number of quality control procedures to ensure that scanning operations resulted in accurate data capture of the teacher-recorded student responses. Prior to the test administration, AIR verified all of the point

values for each form's answer folder. For each form, two AIR staff members independently verified the possible responses and point values for each item.

After receiving the scanned test data, AIR analysts carefully examined the data file to verify its accuracy. Descriptive statistics were computed to ensure that student case counts on the pre-identification file generally corresponded to the actual counts based on test data at the state, school, and classroom levels. In addition, AIR verified that the total number of items in the data file matched the number of items on the answer folder and in the test booklet and then examined the frequency distributions of item responses to identify potential scoring problems, such as out-of-range values or unused response categories.

For purposes of item analysis and student scoring, respectively, non-response (NR) data were treated in two different ways:

For *item analysis and calibration purposes*, a student had to have at least three scored responses for the testing attempt to be considered valid. For a response to be considered a scored response, the test administrator had to have assigned a numeric score (0–4) to the student's response. If the administrator scored NR for all items in a task, the task was treated as not administered, and NR values were recoded as missing.

For *operational scoring* of student responses and estimation of student proficiency, however, the NR codes were treated as indications that the item was administered and that the student did not possess the content area knowledge and skill to respond. In this case, all NR values were recoded as zeroes and included in the student proficiency estimates. Following this recoding, tests were reexamined to determine the number of scored responses (0–4) in each content area. For operational scoring, a student had to have at least five scored responses of any kind for the assessment to be considered a valid attempt within a content area.

After the accuracy of the data file was verified, classical item analyses and Item Response Theory (IRT) analyses were performed. Several quality control procedures were taken to ensure the accuracy of these analyses.

As an initial step, the program control file was checked by two data analysts to ensure that form layout was correctly specified and that item response values were correct. As a second step, two analysts independently performed all analyses. Results of the parallel analyses were compared for mistakes by using commercially available file comparison software. Last, the analysis results were spot-checked by using other commercially available statistical software to ensure that the results were consistent across statistical software packages. *These comprehensive quality control steps are highly effective in detecting any issues that might influence the interpretation of the item analysis results.*

Classical Item Analysis

Classical item analysis for the SC-Alt operational and field-test forms was conducted using the AM statistical software (<http://am.air.org>). The item analysis yielded the proportion of students in each response category, the percentage of omitted responses for that item,⁷ and the proportion of students who were unable to respond to the item because of access limitations (where relevant).

⁷ An item was considered omitted if no response was recorded for the item (or the test administrator marked NR on the student score sheet), but the student responded to subsequent items on the task.

Correlations between the item score and the test score were computed using adjusted polyserial correlations. For purposes of calculating item statistics, omitted items were treated as incorrect when there was at least one scored response within the same task (see above). Minimum and maximum point values, average item scores, and adjusted item-total polyserial correlations were calculated for all items.

Test form statistics, such as internal consistency reliability estimates and standard error of measurement statistics, were suppressed at this point because all students were not expected to take all items. Such statistics would be misleading before Rasch scoring was applied. Special marginal reliability analyses used to determine the reliability of the student score estimates are described in a later section of this chapter.

The proportion of students in each score-point category was calculated as defined by the item's scoring guidelines, as well as the proportion of students with blank responses within attempted tasks (i.e., those with at least one scored response). Item difficulty was computed as the mean score on the item across all students taking the form and with a scored response on that item. The average proportion of total points, calculated as the mean score divided by the total number of points possible on the item, serves as an additional measure of item difficulty.

Review of Items Not Meeting the Specified Psychometric Criteria

Classical item analysis provided information about the technical quality of the items; items failing to meet specified psychometric criteria were flagged for subsequent review. During field-testing of ELA and mathematics (spring 2006), science (fall 2006), and social studies (spring 2007), AIR reviewed all flagged items in concert with SCDE to determine whether they were of sufficient psychometric quality. For the 2007 operational forms in ELA, mathematics, and science and for the 2008 operational form in social studies, AIR conducted a statistical review of the items to determine whether any operational items were performing in an unacceptable fashion. For the spring 2009 operational SC-Alt administration, AIR subjected all embedded field-test items in ELA, mathematics, science, and social studies to an item data review. The spring 2010 operational SC-Alt administration contained only operational items in ELA, mathematics, science, and social studies. Items from the 2010 spring independent field test in high school biology were subjected to similar analyses as with previous field tests. The 2011 administration utilized a field-test design with embedded field-test tasks in ELA, mathematics, science, and social studies, plus operational field-test tasks in science.

Item Response Theory Calibration and Linking Test Forms

This section describes AIR's procedures for item calibration using IRT techniques. Item parameters were estimated using the Partial Credit model (Masters, 1982) approach available using Winsteps software. A common item design was used to enable simultaneous calibration and linking across grade-band test forms in each content area. Items were jointly calibrated across grade-bands in a single Winsteps run for each content area. This calibration approach put the item parameters of all grade-band test forms within a content area on the same scale.

For 2011, the results reported on the vertical scale appear in Exhibit 5.1 and Exhibit 5.2. It is interesting to note that the mean scores show a general upward trend from elementary to middle school levels. This indicates that a vertical scale is a useful way to describe the results of this population of students. The exception is the mean mathematics scale score in high school, which

falls seven scale score points below the middle school mean. In addition, in almost every grade-band, a few students were at the floor of the test (minimum scaled score equal to 260), but very few reached the ceiling (maximum scale score equal to 740).

Exhibit 5.1: Scale Score Statistics, by Grade-Band, Overall

Subject	Statistic	Elementary School	Middle School	High School
ELA	N	1485	1326	355
	Mean	498.12	517.06	517.55
	SD	58.95	71.81	69.38
	Min	260	260	260
	Max	663	725	740
Math	N	1485	1320	350
	Mean	499.62	515.17	508.39
	SD	66.40	68.70	65.99
	Min	260	260	260
	Max	708	721	732
Science/Biology	N	1031	914	350
	Mean	504.73	517.68	509.41
	SD	61.89	74.34	100.55
	Min	260	260	260
	Max	672	733	740
Social Studies	N	1024	922	
	Mean	507.93	521.76	
	SD	68.75	74.97	
	Min	260	260	
	Max	701	740	

Exhibit 5.2: Scale Score Statistics, by Grade-Band, by Primary Disability

Subject	Statistic	Elementary School				Middle School				High School			
		Severe	Moderate	Mild	Autism	Severe	Moderate	Mild	Autism	Severe	Moderate	Mild	Autism
ELA	N	135	332	387	397	134	404	364	301	31	132	97	70
	Mean	408.10	481.91	534.29	491.53	427.19	500.87	564.49	504.44	410.35	511.77	564.37	497.71
	SD	62.95	43.10	39.38	45.38	65.39	48.34	56.11	58.55	81.12	50.09	39.66	68.63
	Min	260	260	409	334	260	260	428	260	260	399	460	260
	Max	548	577	663	630	574	717	725	717	526	740	713	630
Math	N	135	332	387	397	134	403	362	301	31	130	95	69
	Mean	399.73	480.59	540.25	494.41	419.39	499.98	559.48	510.07	407.19	503.36	551.67	493.84
	SD	66.65	46.25	50.42	49.35	69.26	44.96	52.49	55.10	87.02	43.23	49.44	59.99
	Min	260	260	411	335	260	267	431	353	260	386	439	260
	Max	510	580	708	653	525	663	721	718	574	732	732	613
Science/Biology	N	87	226	276	277	92	286	244	215	31	129	96	69
	Mean	404.49	491.28	539.8	499.4	421.54	502.44	566.55	504.10	376.48	495.78	581.14	472.29
	SD	72.38	48.20	40.58	43.84	67.55	48.19	56.75	67.63	104.11	71.11	76.55	90.07
	Min	260	260	424	359	260	260	420	289	260	298	260	260
	Max	525	599	672	636	572	664	733	727	582	740	740	740
Social Studies	N	88	230	282	270	95	269	263	204				
	Mean	398.00	487.77	549.83	500.74	423.17	509.29	569.14	505.74				
	SD	72.30	60.01	42.56	52.51	69.39	53.22	54.9	58.37				
	Min	260	260	419	364	260	260	428	260				
	Max	506	701	701	663	540	690	740	736				

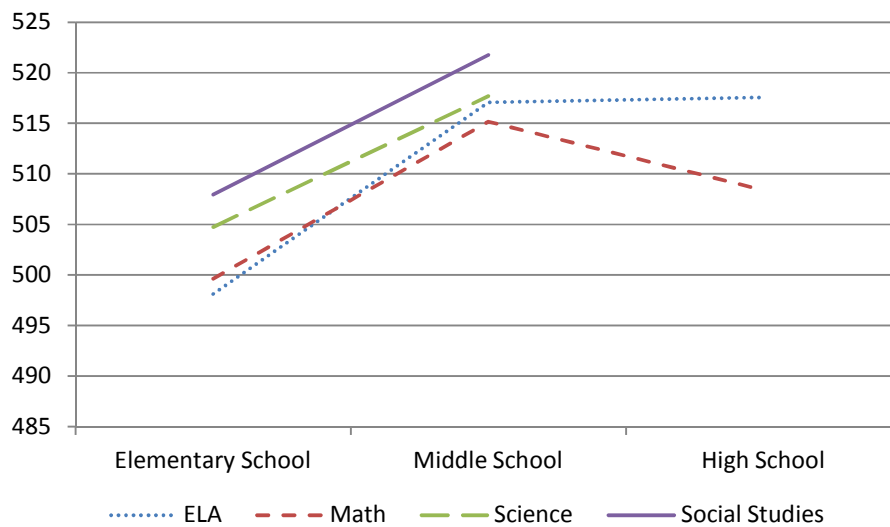
The South Carolina alternate assessment implements vertical scaling of the assessments, permitting the measurement of student progress on the state content standards over time. Such a scale can provide educators and parents with useful information for monitoring student performance as students move through grades over time.

The development of this scale required the use of a common item linking design. In a common item design, *linking item tasks* appear on two adjacent grade-band forms of the assessment. These linking items allow for the grade-band scales to be connected, thus establishing the vertical scale.

There are at least two features of this linking design and the use of grade-level forms that warrant clarification. First, the linking items are the same (i.e., they are the same items) between two grade-band forms. They are connected to grade-specific standards in the higher grade as well as in the lower grade. As a result, students are not exposed to off-grade-level content since the common items serve a dual purpose in measuring content in both grade-bands. Second, even though some items on each grade-band form are administered for two or three years before replacement, it is not likely that exposure of the items to the students would trigger responses based on the recollection of any item's administration the previous year. As students grow academically, their starting task will likely change each year. New starting tasks mean that a portion of the items any student receives will be unique each year. New task development allowed for field-test tasks to be placed on the 2011 form. These field-test tasks will then be placed on the 2012 form as operational tasks, provided that enough of their items passed the 2011 item data review.

The linking design was changed for the 2011 forms and further changes in the use of linking tasks will be made in the 2012 and future forms. For the 2011 forms, in the area of science, biology was added as an operational assessment at the high school level without linking to the earlier grades. Vertical linking tasks were also discontinued for the elementary and middle school science forms. A linking task design was maintained in ELA, mathematics, and social studies. For the 2012 forms and beyond, vertical linking tasks will be used only for ELA and mathematics, and the scores of these linking tasks will contribute operationally to student performance scores at only one grade-band.

A graph of the overall pattern of performance for 2011 on the vertical scale is shown in Exhibit 5.3. Again, there is a general upward trend from elementary to middle school grade-bands in each of the four subject areas. This graph shows that the vertical scale in the SC-Alt was successful at capturing growth across grade-bands.

Exhibit 5.3: Overall Pattern of Performance on the Vertical Scale

Using Item Responses to Estimate Student Proficiency

This section describes the estimation of student proficiency for the SC-Alt operational administration of English language arts, mathematics, and science/biology assessments for elementary, middle, and high school; social studies assessments for elementary and middle school are also reported. The section describes the estimation procedures used to determine student proficiency based on the items administered, the transformation of proficiency estimates on the Rasch theta scale into scale scores, and finally the relation of achievement estimation to reliability estimation.

Student proficiency scores were estimated using a maximum-likelihood approach based on the scored items for each student.⁸ This method calculates the theta score that maximizes the likelihood function of the given item responses for each student. Comparable scale-score estimates from these different item responses were achieved through the measurement-invariance property of IRT ability estimates, even when students were exposed to different ranges of items.

Under the Rasch-based IRT model, there is a one-to-one correspondence between the estimated theta score and the total raw score for a specific set of items. However, in the SC-Alt, each student can take different sets of items. Using the pattern scoring method for calculating theta scores, we ensured that (a) two students who took the same items and achieved the same item scores were assigned the same theta score, and (b) students who took more difficult items were assigned higher theta scores than students with the same raw scores who took less difficult items. Thus, the scoring method took into account both the number of raw score points the student achieved and the difficulties of the items the student responded to. This scoring process was performed separately for each content area.

⁸ The first step in this process was to rescore student responses consistent with the operational scoring method described under the “Data Preparation and Quality Control” section.

Once theta values had been estimated for each student, AIR converted the theta estimates to scale scores using a scale metric determined by SCDE in consultation with AIR. The SC-Alt assessments were scaled to have a mean of 500 and a standard deviation of 80 on the vertical scale for the grade-band 6–8 assessment. The grade-band 3–5 and grade 10 assessment means and standard deviations were calculated in relation to the grade-band 6–8 mean and standard deviation. This was done by performing a linear transformation of the Rasch theta scale for each content area, fixing the mean of the middle school test form scale at 500, and multiplying the student’s theta deviation score by 80 as shown in the formula below:

$$y_{ijk}^* = 500 + \left(\frac{\hat{\theta}_{ijk} - \hat{\mu}_k}{\hat{\sigma}_k} \right) * 80$$

where

i indexes student;

j indexes grade-band;

k indexes content area;

y_{ijk}^* is the scale score for student i in grade-band j and content area k , given estimated ability, θ_{ijk} ;

$\hat{\mu}_k$ is the content-area-specific mean for the middle school test form; and

$\hat{\sigma}_k$ is the content-area-specific standard deviation for the middle school test form.

SCDE also decided to truncate the scale score ranges so that the lowest possible scale score was 260 and the highest possible scale score was 740. Student scale-score estimates were truncated to the smallest whole integer (e.g., an estimated scale score of 440.60 would become 440). Additionally, scale scores were calculated and checked using a method similar to the process for total raw data.

Once scoring was completed, it was possible to estimate the internal consistency score reliability of the grade-band assessments by estimating the marginal measurement error across students. These estimates produced different standard errors for each student, depending on the items they were given and their level of performance on those items. This value was used to determine the score reliability as the proportion of true score variance to observed score variance. We estimated this value within each content area (a) across the entire theta scale, (b) across grade-band forms, and (c) for each starting point within a grade-band.

Test Score Reliability

This section provides the marginal reliability for each grade-band, content area, and groups of students beginning at each starting task determined by the SPQ for the spring 2011 administration.

Classical test theory-based reliability indices, such as Cronbach's alpha, were not appropriate for the SC-Alt because the length of the test and the subset of items differed for each student. The reliability coefficient for the SC-Alt was, therefore, calculated as the *marginal reliability* (Sireci, Thissen, & Wainer, 1991), which is equivalent in interpretation to classical internal consistency estimates of reliability.

First we determined the marginal measurement error variance, $\bar{\sigma}_e^2$, across all examinees with a score strictly between the score limits of 260 and 740:

$$\bar{\sigma}_e^2 = \int \sigma_e^2 p(\theta) d\theta = \frac{\sum \sigma_e^2}{N},$$

where σ_e^2 is the square of the standard error of student ability estimate, $\hat{\theta}$. Thus, the marginal measurement error variance could be estimated as the average of squared standard error of $\hat{\theta}$.

Then we estimated the marginal reliability as

$$\bar{\rho} = \frac{\hat{\sigma}_\theta^2 - \bar{\sigma}_e^2}{\hat{\sigma}_\theta^2},$$

where $\hat{\sigma}_\theta^2$ is the variance of observed θ estimates.

The marginal reliability estimate, $\bar{\rho}$, can be interpreted similarly to classical reliability indices such as Cronbach's alpha. Estimates of the marginal reliability for the test forms corresponding to the three SC-Alt grade-band assessments can be seen in Exhibit 5.4.

Exhibit 5.4: Marginal Reliability and Standard Error of Measurement by Grade-Band and Subject

Subject	Grade-Band	Elementary	Middle	High
English Language Arts	N	1473	1314	349
	Reliability	0.93	0.90	0.93
	$\bar{\sigma}_e^*$	14.37	21.72	16.61
Math	N	1465	1305	344
	Reliability	0.92	0.91	0.91
	$\bar{\sigma}_e^*$	17.16	19.15	16.96
Science/Biology	N	1019	908	326
	Reliability	0.90	0.89	0.87
	$\bar{\sigma}_e^*$	17.72	23.98	29.53
Social Studies	N	1007	895	
	Reliability	0.92	0.91	
	$\bar{\sigma}_e^*$	16.92	19.17	

The marginal reliability estimates for ELA, mathematics, and social studies met or exceeded 0.90 in each grade-band form; and the marginal reliability estimates for science and biology met or exceeded 0.87. The reliability estimates of all four content areas fall into the range of reliability coefficients found with large-scale assessments (Rudner & Schafer, 2001) and meet the reliability requirements for assessments used for the purposes for which the SC-Alt was designed.

In addition to the marginal reliability estimates, Exhibit 5.4 also displays the marginal standard errors of measurement for each subject and grade-band, labeled $\bar{\sigma}_e^*$. These marginal standard errors of measurement range between 14 and 30 scale score units, placing the SEM at approximately from a quarter to just less than a third of a standard deviation of the content area and grade-band.

Appendix G shows the marginal reliability estimates broken out by groups of students beginning at each starting task, as determined by the SPQ. The reliability coefficients in Exhibits G-1 through G-4 are generally somewhat attenuated compared to those in Exhibit 5.4, due to the reduction in variance of scale scores grouped by starting task.

Appendix G also displays the marginal reliability coefficients for each subject and grade-band, broken down by gender (Exhibit G-5) and by the major ethnicity subgroups (African American vs. White; Exhibit G-6). The coefficients in Exhibits G-5 and G-6 range in the high eighties and low- to mid-nineties, indicating acceptable reliabilities for these demographic subpopulations.

Classification Accuracy

This section describes the extent to which student achievement-level classifications were accurate across students. Classification accuracy was estimated for each cut score as the average probability of correct achievement-level assignments across all examinees (assignments above or below the cut score), given each examinee's estimated proficiency score, θ_i :

$$CA_K = \frac{\sum_{i=1}^{N_{k \geq K}} P(\theta_i > \theta_K^* | \theta_i, k_i \geq K) + \sum_{i=1}^{N_{k < K}} [1 - P(\theta_i > \theta_K^* | \theta_i, k_i < K)]}{N},$$

where

θ_i is the proficiency (i.e., theta) of student i ;

k_i is the assigned performance level of student i ;

θ_K^* is the cut score for the performance level K on the theta scale; and

N is the sum of the number of students at or above the cut score, $N_{k \geq K}$, and the number of students below the cut score, $N_{k < K}$, or simply the total number of students.

Thus, $P(\theta_i > \theta_K^* | \theta_i, k_i \geq K)$ is the probability that a student with θ_i assigned to achievement level k_i is above the cut score, θ_K^* . The classification accuracy is the expected rate of correct classification probability, ranging from 0 to 1, where higher values indicate superior classification consistency. Exhibit 5.5 shows the classification accuracy by content areas, achievement levels, and grade-bands.

Exhibit 5.5: Classification Accuracy

Subject	Achievement Level	Elementary School	Middle School	High School	Overall
English Language Arts	Level 2	0.991	0.988	0.986	0.989
	Level 3	0.937	0.940	0.944	0.939
	Level 4	0.946	0.940	0.960	0.945
Mathematics	Level 2	0.986	0.982	0.987	0.984
	Level 3	0.941	0.946	0.921	0.941
	Level 4	0.936	0.940	0.933	0.937
Science	Level 2	0.981	0.975		0.979
	Level 3	0.942	0.932		0.938
	Level 4	0.910	0.931		0.920
Biology	Level 2			0.960	0.960
	Level 3			0.903	0.903
	Level 4			0.905	0.905
Social Studies	Level 2	0.981	0.976		0.979
	Level 3	0.936	0.932		0.934
	Level 4	0.924	0.916		0.920

For example, according to the estimates in Exhibit 5.5, for the grade-band 3–5 English language arts assessment, 99% of students were correctly classified at Level 2 or higher (versus at Level 1) and 94% of students at Level 3 or above (vs. at Levels 1 or 2 combined). All students in all grade-bands had a probability greater than 0.90 of being classified accurately as proficient (i.e., as Level 3 or higher) vs. not proficient. The highest classification accuracy for Level 3 or higher was found for ELA and mathematics, the lowest for biology.

These results indicate that the measurement errors at the performance-level cut points for ELA, mathematics, science, and social studies (Exhibits 4.1–4.4) are small compared to the overall variance of student performance. On the other hand, the measurement error at the Level 3 cut point for biology (Exhibit 4.5) appears somewhat elevated.

The calculation of the probability of the correct performance level for students is described in the following section.

Calculation of the Probability of Being Classified Above a Cut Score Given the Student's Theta Score

For each student we can compute the likelihood of theta $L(\theta | \mathbf{z}, \mathbf{b})$. Suppose that the prior of the theta distribution is $f(\theta)$. Then, using Bayes' rule, we have

$$f(\theta | \mathbf{z}, \mathbf{b}) \propto f(\theta)L(\theta | \mathbf{z}, \mathbf{b}),$$

where $L(\theta | \mathbf{z}, \mathbf{b})$ is the likelihood of theta given the response \mathbf{z} and item parameters \mathbf{b} ; hence, the probability at and above cut is

$$P = \frac{\int_{\theta \geq \theta_{\text{cut}}} f(\theta) L(\theta | \mathbf{z}, \mathbf{b}) d\theta}{\int f(\theta) L(\theta | \mathbf{z}, \mathbf{b}) d\theta},$$

where $f(\theta)$ can take different distribution such as normal, or uniform, depending on our prior belief.

Calculation of the Likelihood of Theta Given Item Scores z and Step Parameters b

For the Rasch model, we have

$$\begin{aligned} L(\theta | \mathbf{z}, \mathbf{b}) &= \prod_{i \in MC} \left(\frac{\text{Exp}(z_i \theta - b_i z_i)}{1 + \text{Exp}(\theta - b_i)} \right) \prod_{i \in CR} \left(\frac{\text{Exp}(z_i \theta - \sum_{k=1}^{z_i} b_k)}{1 + \sum_{i=1}^{K_i} \text{Exp}(\sum_{k=1}^i (\theta - b_k))} \right) \\ &\propto \text{Exp}(r \theta) \prod_{i \in MC} \left(\frac{1}{1 + \text{Exp}(\theta - b_i)} \right) \prod_{i \in CR} \left(\frac{1}{1 + \sum_{i=1}^{K_i} \text{Exp}(\sum_{k=1}^i (\theta - b_k))} \right) \end{aligned}$$

where K_i is the maximum score for item i when this item is a CR item. It can be noted that the calculation above depends on total raw score r only when using the attempted items.

Chapter 6: Score Reports

This chapter describes the method used for reporting scores on the SC-Alt for the spring 2011 administration. An Individual Score Report (ISR) is included in Appendix H as an example of the highly detailed and diagnostic nature of the reports. This chapter gives a brief overview of how scores on the SC-Alt are reported; a more detailed description is available in a separate *Score Reports User's Guide*.

The SC-Alt has three types of score reports: the ISR, or family report; school reports; and district reports. Each report conveys specific information to its target audience. The reports are designed to be easily used by parents and educators. Of particular note, the reports include in-depth information about what students know and can do relative to the South Carolina academic content standards and to the performance levels.

The ISR provides specific performance feedback for each student across these content areas: English language arts, mathematics, science, and social studies in grade-bands 3–5 and 6–8; and English language arts, mathematics, and biology in grade 10. Within each content area, a graphic bar highlights the student's performance level along the proficiency scale. Each performance level is described in broad, easy-to-understand content terms. Further descriptions of what a student knows and can do are tailored and printed for each obtained performance level. For example, if a student is classified as Level 3 in mathematics, the following message is printed: "Students who score at Level 3 should be able to add and subtract simple numbers, count and compare objects in a group, compare objects by color, size, or shape, identify three-dimensional shapes, and read information in a graph." (Note: Scale scores were added to the ISR starting with the spring 2008 reporting cycle.)

Specific activities, based on each student's performance level for each content area, are presented for the family to do at home to help ensure positive academic growth in the content area.

The school report provides a summary of the performance of each student in the school. The alphabetical list of students contains basic demographic information and test form administered, in addition to achievement data. A scale score and achievement level are listed for each student for each content area. A school summary shows the number of students scoring at each performance level.

Three district-level reports are issued. The district roster summary report displays the roster of the district's tested students along with their demographic information, their scale scores and performance levels for each content area, and type of test form. The district summary by test form report presents a roster of schools in which students were tested, identifying the test form and giving the number of students tested in each content area and the percentages achieved in each performance level by content area. The total number of students tested with each form and their performance-level distributions by content area are listed at the bottom of the report. The district demographic summary report shows the number of students tested and the distributions of performance levels in all content areas, disaggregated by gender, ethnicity, lunch program, migrant status, and ESL status.

The separate *Score Reports User's Guide* has more specific information on how to interpret student scores and score reports and how to relate academic growth as measured by the SC-Alt to

classroom curricula and activities. The guide has been widely distributed throughout South Carolina.

Chapter 7: Student Performance Data from the Spring 2011 Administration

Performance data from the spring 2011 administration are presented in this chapter. This was the fifth operational administration of the SC-Alt ELA, mathematics, and science assessments; the fourth operational administration of the SC-Alt social studies assessment; and the first operational administration of the SC-Alt high school biology assessment.

A total of 3,184 students from 84 school districts and 563 schools were tested with the SC-Alt in spring 2011. The total number of tested students with one or more valid content area scores was 1,486 for the elementary form, 1,326 for the middle school form, and 355 for the high school form.

Nearly one-half of the participating school districts (37; 44%) tested 15 or fewer students; 27 districts (32%) tested 16 to 50 students; and 20 districts (24%) tested more than 50 students each. Eight districts tested more than 100 students; the greatest number of students tested in one district was 256.

Of the 563 schools testing SC-Alt students, 364 (65%) tested five or fewer students; 131 (23%) tested six to 10 students; 56 (10%) tested 11 to 20 students; and 12 schools (2%) tested 21 or more. Only two schools tested more than 50 students each (64 and 73 students).

The elementary school form was developed to be administered to students who are 8, 9, or 10 years old at the beginning of the school year, which are the ages typical of students enrolled in grades 3, 4, and 5. The middle school form was developed for students who are 11, 12, and 13 (typical of students enrolled in grades 6, 7, and 8), and the high school form was developed for students age 15 (typical age of students in grade 10).

Students tested with the elementary and middle school forms with reported ages outside the specified age ranges were either erroneously assigned to the forms by the test administrator or, in some cases, took the test as a result of birth date coding errors on the data files. The number of students reported outside the expected ages for the elementary and middle school forms is less than 1% for each content area. Students reported as having been tested on the high school form with ages below 15 were tested because of errors in form assignment or birth date coding. Students older than 15 (e.g., 16) may be assessed with the high school form if they have not been assessed at the high school level previously.

The performance of students by grade-band form, age, and demographic group for the ELA, mathematics, science, and social studies content areas is presented in Appendix I.

Chapter 8: Validity

Content Validity

One source of evidence for the content validity of the South Carolina Alternate Assessment was obtained through independent alignment studies. The University of North Carolina at Charlotte (UNCC) conducted studies of the alignment of (a) Assessment Standards and Measurement Guidelines (ASMGs) to grade-level curriculum standards and (b) SC-Alt items to the ASMGs that they targeted. This was a pilot study conducted by Flowers, Browder, Wakeman, and Karvonen with UNCC through the National Alternate Assessment Center (NAAC). (South Carolina is a member state of the NAAC.) A second independent study of ELA and mathematics was completed by the South Carolina Education Oversight Committee (EOC; 2008a) as required by the state Education Accountability Act of 1998 (EAA). The EOC approved the ELA and mathematics content areas on February 28, 2008. The UNCC alignment study results for the English language arts and mathematics assessments are reported in detail in Flowers, Browder, Wakeman, and Karvonen (2006a). The results of the alignment studies for the ELA and mathematics assessments indicate that

The state has evidence supporting alignment for its measurement guidelines and alternate assessment based on all seven criteria. We conclude that overall this is an alternate assessment system that links to the grade level content. Some areas for consideration in further development of the system are noted related to balance of content. (p. 7)

The alignment study results for the science assessment are reported in detail in Flowers, Browder, Wakeman, and Karvonen (2006b) and in an addendum dated December 21, 2007. The results of the alignment study for the science assessment indicate that

The strength of the South Carolina science Alternate Assessment was that nearly all of the content was academic science content (98%). This is especially notable given that the alternate assessment tasks included items accessible to students at all symbolic levels. In contrast, the degree of alignment of AA tasks/items to grade-level standards was lower than those found in the alignment of ELA and mathematics. This difference could be due to the fact that the state's science grade-level standards changed during the development of the science AA. Another challenge was that the state had linked its alternate assessment tasks to the state standards and not directly to the measurement guidelines, creating a tough challenge to demonstrating alignment.... Our work with other states suggests that science may typically be the area rated as having the weakest alignment. (p. 4)

SCDE reviewed the initial science alignment study and determined that one source of some misalignment had resulted from the linking of some items to multiple standards and indicators in the alignment document provided by SCDE. During the Science Content Review Committee meeting, some members recommended adding additional indicators to align to some items. The intent of these recommendations focused more on instruction and demonstrating that instruction could include multiple standards and indicators. However, the alignment study team considered only the first two standards aligned to each item. In some cases, the first two standards were not necessarily the most appropriate. SCDE prioritized the standards and indicators and resubmitted the documentation for an additional study. From this review, completed December 21, 2007

(Flowers, Browder, Wakeman, & Karvonen, 2007), 163 of 173 items were rated as academic. Of the 10 items listed as nonacademic, six were rated as foundational (p. 1). SCDE is currently addressing the items that were rated as having no content centrality by developing replacement items for new forms.

At the time of the alignment study for ELA and mathematics by Dr. Flowers and colleagues, the design of the SC-Alt was envisioned as a single assessment across grade levels. This design changed to a grade-band assessment following the study; however, the information provided from the alignment study was used to identify items with alignment difficulty, and these items were omitted from the operational grade-band test forms. Information from the review along with teacher comments were also used during item data review as part of the decision-making process regarding inclusion of items in the assessment.

A second independent review of the alignment of the science assessment was conducted by the Education Oversight Committee (EOC; 2008b). The EOC approved the elementary and middle school science alternate assessment on August 12, 2008. The EOC alignment findings were based on the review of two sets of studies of the SC-Alt:

- Studies of the alignment between the SC-Alt science assessment and the state academic standards conducted by University of North Carolina-Charlotte and Western Carolina University professors of curriculum and special education, in cooperation with the South Carolina State Department of Education (SCDE) and the National Alternate Assessment Center (Flowers, Browder, Wakeman, & Karvonen, 2006a, 2006b, 2007)
- A technical review of the task and item data from the 2007 test administration conducted by a professor of educational research and assessment at the University of South Carolina

Copies of the reports of the EOC reviews and findings are available in their entirety from the SCDE. Based on this review, the EOC identified a number of strengths of the SC-Alt science assessment that were noted in the final report:

- The assessment provides accountability and information for instructional improvement for students with significant cognitive disabilities who would not otherwise be assessed in the state testing programs, even with test accommodations and modifications.
- The assessment is intended to be aligned with the same grade-level academic standards as for all students, although at levels of complexity appropriate for the diversity of cognitive functioning observed among students with significant cognitive disabilities.
- The assessment format allows each student to respond to the items using the communication modes the student uses during instruction, such as oral response, pointing, eye gaze, a response card, sign language, or an augmentative communication device.
- The procedures for placing the student at the appropriate level for beginning each assessment reduces student fatigue and maximizes the student's opportunities to show his or her highest performance;
- The items in the assessment have a wide range of difficulty, and the test is moderately able to discriminate between high and low levels of performance.

The EOC report noted that while 96% of the items were found to be aligned to science inquiry standard indicators, the alignment of the items to content standards was 78%, falling short of an expectation for successful alignment of 90% set by the original evaluators. The EOC recommended that the SCDE review the alignment of the SC-Alt science items to the grade-level standards and identify items needing revision or replacement.

The SCDE and its contractor, the American Institutes for Research (AIR), reviewed the alignment and the ASMGs and established priorities for development of tasks to fill identified gaps. During 2008, SCDE and AIR developed five new tasks consisting of 32 items to be used to replace poorly aligned items and improve content coverage in science. Three tasks were developed for the elementary science form, and two tasks were developed for the middle school form based on the findings of the alignment study. The high school physical science test was replaced by a high school biology assessment in spring 2010.

An independent review of the alignment of the new items by the Center for Research on Education (2009a) found that 98% of the new items were aligned to grade-level content standard indicators. Copies of the report of the alignment reviews and findings are available in their entirety from the SCDE.

A follow-up alignment study of the high school ELA and mathematics assessments and biology field-test items was conducted by the Center for Research on Education in October 2009, using the same procedures that were used for the elementary and middle school alignment studies in December 2006 and January 2007. Almost all (94% to 96%) of the items were rated as academic. This percentage exceeds the value typically found in alternate assessments (90%) according to the reviewers. The alignment study results are reported in detail in *High School Alternate Assessment Alignment Report to the South Carolina State Department of Education* (Center for Research on Education, 2009b).

Convergent and Discriminant Validity

According to Critical Element 4.1(e) of the federal peer review and Standard 1.14 of the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999), it is desirable, if not necessary, to provide evidence of convergent and discriminant validity. One common method for examining this aspect of validity is with a multitrait-multimethod matrix (MTMM) (Campbell & Fiske, 1959).

Campbell and Fiske (1959) proposed the MTMM matrix design as a tool for the study of convergent and discriminant validity in psychological measurement. The MTMM matrix employs a crossed factorial measurement design of traits and methods to reveal these types of validity in comparison:

- Large correlations on validity diagonals (i.e., same trait and different methods) indicate convergent validity.
- Low correlations in the heterotrait-monomethod blocks indicate discriminant validity and the absence of method effects.
- Low correlations in the heterotrait-heteromethod blocks also indicate discriminant validity.

Selection of Traits and Methods

The student's abilities in each of the subjects—ELA, mathematics, science/biology and social studies—make up the four traits for the MTMM study. Two methods are considered for assessing these traits: the Student Placement Questionnaire (SPQ), as a structured teacher rating of student's attainment, and the SC-Alt scale score, as an IRT-based indicator of the student's performance in each subject or trait. In other words, the two methods contrast test scores of student performance with expert (or teacher) ratings. With four traits and two methods, the MTMM correlation matrix is of order 8. Note that the high school assessment does not include a social studies component; therefore, the MTMM matrix for high school has only six rows and columns.

Results

MTMM matrices were computed separately for each grade-band. The results are given in Exhibits 8.1–8.3. Pearson correlations are used, with pairwise deletion of missing data. For each matrix, the minimum pairwise sample size is indicated. *P*-values of individual correlation coefficients are not reported since *all* correlations are significant ($p < 0.05$).

Exhibit 8.1: MTMM Matrix, Scale Scores with SPQ Scores, Elementary School

	Subject	IRT Scale Scores				SPQ Scores			
		ELA	Math	Science	Social Studies	ELA	Math	Science	Social Studies
IRT Scale Scores	ELA	1.000							
	Math	0.892	1.000						
	Science	0.888	0.894	1.000					
	Social Studies	0.900	0.893	0.894	1.000				
SPQ Scores	ELA	0.774	0.760	0.717	0.764	1.000			
	Math	0.764	0.772	0.715	0.746	0.916	1.000		
	Science	0.746	0.734	0.713	0.725	0.876	0.895	1.000	
	Social Studies	0.759	0.741	0.704	0.774	0.906	0.896	0.936	1.000
Minimum pairwise N: 576									

Exhibit 8.2: MTMM Matrix, Scale Scores and SPQ Scores, Middle School

		IRT Scale Scores				SPQ Scores			
	Subject	ELA	Math	Science	Social Studies	ELA	Math	Science	Social Studies
IRT Scale Scores	ELA	1.000							
	Math	<i>0.888</i>	1.000						
	Science	<i>0.877</i>	<i>0.883</i>	1.000					
	Social Studies	<i>0.887</i>	<i>0.869</i>	<i>0.893</i>	1.000				
SPQ Scores	ELA	0.750	0.766	0.745	0.764	1.000			
	Math	0.763	0.791	0.760	0.766	<i>0.918</i>	1.000		
	Science	0.744	0.760	0.762	0.748	<i>0.887</i>	<i>0.912</i>	1.000	
	Social Studies	0.760	0.774	0.756	0.773	<i>0.908</i>	<i>0.907</i>	<i>0.924</i>	1.000
Minimum pairwise N: 511									

Exhibit 8.3: MTMM Matrix, Scales Scores and SPQ Scores, High School

		IRT Scale Scores			SPQ Scores		
	Subject	ELA	Math	Biology	ELA	Math	Biology
IRT Scale Scores	ELA	1.000					
	Math	<i>0.888</i>	1.000				
	Biology	<i>0.867</i>	<i>0.847</i>	1.000			
SPQ Scores	ELA	0.754	0.692	0.712	1.000		
	Math	0.748	0.712	0.720	<i>0.912</i>	1.000	
	Biology	0.531	0.523	0.564	<i>0.608</i>	<i>0.692</i>	1.000
Minimum pairwise N: 316							

In each MTMM table, the *convergent validity coefficients* (correlations between measurements of the same trait using different methods) are marked in bold. Most of the convergent validity coefficients range from 0.71 to 0.79 and certainly fall into an acceptable range. These high correlations demonstrate evidence for the validity of the SPQ; the three exhibits indicate that the SPQ and the actual test are essentially measuring the same trait and that the SPQ is a good indicator of performance on the test. The one exception is the high school biology assessment for which the SPQ and the IRT scale scores correlate at 0.564; the reason appears to lie with the biology SPQ score, as it shows noticeably low correlations with *all* the other variables, whether SPQ scores or IRT scale scores. The biology SPQ operates somewhat differently from the other tests; this may be an interesting target for further investigation.

The *heterotrait-monomethod coefficients* in the monomethod triangles (correlations between measurements of different traits using the same method) are set in italics. These correlation coefficients range between 0.84 and 0.90 for IRT scale scores and between 0.61 and 0.94 for SPQ scores. The high overall range of these correlations (with the exception of lower correlations involving SPQ scores in biology) indicates the presence of method variance.

However, this is to be expected because the SPQ was not developed to measure the trait; instead, it only indicates the starting task on the test for measuring the trait. Such a result of high correlations in the monomethod triangles is not uncommon in MTMM studies (Fiske, 1995), and specific conditions offer themselves as causes for the present scenario. First, the different scale types—number-correct rating scales versus IRT scales of behavioral tests—are in themselves a source of method variation; second, the SPQ’s “can do” questions draw on the teacher’s memory of a student’s possible performance over the long term and are apt to differ in quality and veracity; and third, the IRT scale scores for the three subjects reflect the student’s performance in the testing situation and are subject to the student’s condition on the testing day.

The *heterotrait-heteromethod coefficients* appear in the tables in regular type. These correlation coefficients fall in the same range as the convergent validity coefficients, with values from 0.52 to 0.77 (again, with the lowest values involving the SPQ scores for high school biology). To confirm discriminant validity, the heterotrait-heteromethod correlations should be smaller than the convergent validity coefficients. However, these MTMM matrices support the notion that all the traits except biology vary essentially along just a single dimension. Because the population of alternate assessment students is so *very* heterogeneous, the students’ general level of cognitive functioning dominates the relationship among their scale scores.

Validity of the Student Placement Questionnaire (SPQ)

AIR reviewed item data from the 2011 administration regarding the agreement between SPQ recommended start points and the final observed start points. The purpose of the study was to determine the effectiveness of the SPQ in identifying the most appropriate starting task.

Administration of the SC-Alt uses the SPQ as a pre-assessment instrument to determine the most appropriate starting point in the assessment. The SPQ requires the teacher to evaluate the student on 12 to 15 “can do” statements addressing the student’s skills and knowledge in each content area on the basis of the teacher’s prior instructional knowledge of the student. A total score computed from the teacher’s SPQ responses indicates the initial starting task for the assessment.

The instructions for using the SPQ require teachers to adjust the starting point below the SPQ recommended start point when the student is not successful on the first administered task. Alternatively, after reviewing the assessment, some teachers may have judged that a student needed to start at a higher level than recommended by the SPQ.

A summary of the results of the agreement between the SPQ recommended start points and the observed start points for each content area and grade-band form is presented in Exhibits 8.4–8.7.⁹ These results indicate that the agreement between the SPQ recommended start point and the observed start point was 98% for ELA, 98% for mathematics, 99% for science and biology, and 98% for social studies administrations. Use of the SPQ pre-assessment score is only the first step in the procedure used by the test administrator in determining where the student should start the assessment. Since the test administrator is required to make adjustments based on the student’s success on the first task, and these adjustments are reflected in the agreement rates, the SPQ appears to be working very effectively for targeting the first task to begin the assessment process.

⁹ Data of students with missing SPQ scores were excluded from these Exhibits.

Exhibit 8.4: Agreement Between SPQ and Observed Start Points by SPQ Recommended Starting Tasks—ELA

Observed Start Task	Elementary School				Middle School				High School				Overall	
					Recommended Starting Task									
	1	3	7	Total	1	3	7	Total		1	3	7		Total
Starting task consistent with SPQ	99.7%	93.9%	98.1%	97.6%	99.6%	97.4%	98.2%	98.4%		100.0%	94.1%	97.5%	97.3%	97.9%
Lower start task than recommended	0.0%	3.0%	1.9%	1.7%	0.0%	0.9%	1.5%	1.1%		0.0%	5.9%	2.5%	2.7%	1.5%
Higher start task than recommended	0.3%	1.7%	0.0%	0.4%	0.4%	0.9%	0.0%	0.2%		0.0%	0.0%	0.0%	0.0%	0.3%
Nonstandard start task	0.0%	1.4%	0.0%	0.3%	0.0%	0.9%	0.3%	0.3%		0.0%	0.0%	0.0%	0.0%	0.3%
No valid test items; no starting task	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%		0.0%	0.0%	0.0%	0.0%	0.0%
Incomplete SPQ	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%		0.0%	0.0%	0.0%	0.0%	0.0%
Inconsistent with SPQ	0.3%	6.1%	1.9%	2.4%	0.4%	2.6%	1.8%	1.6%		0.0%	5.9%	2.5%	2.7%	2.1%
ELA Total (N)	356	296	790	1442	268	233	794	1295		67	68	202	337	3074

Exhibit 8.5: Agreement Between SPQ and Observed Start Points by SPQ Recommended Starting Tasks—Math

Observed Start Task	Elementary School				Middle School				High School				Overall	
					Recommended Starting Task									
				Total				Total				Total		
	1	3	7		1	3	7			1	3	7		
Starting task consistent with SPQ	99.4%	94.6%	98.4%	97.7%	100.0%	94.6%	99.2%	98.3%		98.5%	91.6%	95.9%	95.5%	97.7%
Lower start task than recommended	0.0%	3.7%	1.7%	1.8%	0.0%	2.9%	0.5%	1.0%		0.0%	4.2%	4.1%	3.3%	1.6%
Higher start task than recommended	0.6%	1.1%	0.0%	0.4%	0.0%	0.4%	0.0%	0.1%		1.5%	2.8%	0.0%	0.9%	0.3%
Nonstandard start task	0.0%	0.6%	0.0%	0.1%	0.0%	2.2%	0.3%	0.6%		0.0%	1.4%	0.0%	0.3%	0.4%
No valid test items; no starting task	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%		0.0%	0.0%	0.0%	0.0%	0.0%
Incomplete SPQ	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%		0.0%	0.0%	0.0%	0.0%	0.0%
Inconsistent with SPQ	0.6%	5.4%	1.7%	2.3%	0.0%	5.4%	0.8%	1.7%		1.5%	8.5%	4.1%	4.5%	2.3%
Math Total (N)	336	351	729	1416	233	276	759	1268		65	71	195	331	3015

Exhibit 8.6: Agreement Between SPQ and Observed Start Points by SPQ Recommended Starting Tasks—Science/Biology

Observed Start Task	Elementary School				Middle School				High School				Overall
					Recommended Starting Task								
	1	3	7	Total	1	3	7	Total		1	3	6	
Starting task consistent with SPQ	99.3%	96.4%	98.7%	98.4%	99.5%	98.4%	99.2%	99.1%	100.0%	94.1%	94.1%	97.8%	98.6%
Lower start task than recommended	0.0%	1.8%	1.3%	1.0%	0.0%	0.5%	0.4%	0.3%	0.0%	2.9%	2.0%	0.9%	0.7%
Higher start task than recommended	0.7%	1.8%	0.0%	0.6%	0.0%	0.5%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.3%
Nonstandard start task	0.0%	0.0%	0.0%	0.0%	0.5%	0.5%	0.4%	0.5%	0.0%	2.9%	3.9%	1.2%	0.4%
No valid test items; no starting task	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Incomplete SPQ	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Inconsistent with SPQ	0.7%	3.6%	1.3%	1.6%	0.5%	1.6%	0.8%	0.9%	0.0%	5.9%	5.9%	2.2%	1.4%
Science/Biology Total (N)	284	223	478	985	205	189	480	874	203	68	51	322	2181

Exhibit 8.7: Agreement Between SPQ and Observed Start Points by SPQ Recommended Starting Tasks --- Social Studies

Observed Start Task	Elementary School				Middle School				Overall
				Total				Total	
	1	3	7		1	3	7		
Starting task consistent with SPQ	98.6%	97.7%	99.1%	98.8%	100.0%	94.2%	98.5%	98.0%	98.4%
Lower start task than recommended	0.0%	1.7%	0.7%	0.7%	0.0%	3.3%	1.0%	1.3%	1.0%
Higher start task than recommended	1.4%	0.0%	0.0%	0.3%	0.0%	1.3%	0.0%	0.2%	0.3%
Nonstandard start task	0.0%	0.6%	0.2%	0.2%	0.0%	1.3%	0.5%	0.6%	0.4%
No valid test items; no starting task	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Incomplete SPQ	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Inconsistent with SPQ	1.4%	2.3%	0.9%	1.2%	0.0%	5.8%	1.6%	2.1%	1.6%
Social Studies Total (N)	220	177	583	980	141	154	582	877	1857

Start-Stop Analysis

Data from the 2011 SC-Alt assessment were analyzed to address two questions concerning SC-Alt administration procedures and student performance:

1. How many tasks and items were administered to students who were started in the assessment at each of the three start points?
2. What was the achievement level performance of students who were started in the assessment at each of the three start points?

To address these questions, the task start point was identified for each student assessed by the 2011 administration of the SC-Alt assessment for all content areas and grade-band forms. According to each task start point, the number of tasks and items administered and the achievement-level distribution were calculated and summarized.

SC-Alt test administrators were instructed to follow specific procedures concerning the use of the Student Placement Questionnaire (SPQ) to determine task start points, the minimum number of tasks to be administered, and whether to continue the administration through additional tasks until the student is no longer able to respond successfully. These procedures are detailed in Appendix B and in the 2011 *Test Administration Manual*, Appendix N.

Number of Tasks Administered

For English language arts, mathematics, science, and social studies, the minimum number of overall tasks to be administered is seven tasks, when the test administration is started at Task 1, or nine tasks, when the administration begins at either Task 3 or Task 7. For assessments in these subjects, the tasks in positions 5, 8, and 12 are field-test tasks. In other words, the sets of seven or nine overall tasks to be administered translate into six or seven operational tasks, respectively.

For Biology, the minimum number of overall tasks to be administered is six when the student was started at Task 1, or seven tasks when the administration begins at either Task 3 or Task 7.

The actual number of tasks administered to students in the ELA, mathematics, science/biology, and social studies content areas for each form level and task start point are presented in Exhibits 8.8–8.11. Note that these exhibits show only the cases for which the assessment started at one of the three standard starting points. For a few students, however, the assessment was begun at some non-standard starting task. These cases are not included in Exhibits 8.8-8.11.

In general, most students were administered at least the minimum number of tasks; the distribution of actual tasks administered often exceeded the minimum required when students were started at Task 1 or Task 3. In ELA, 99% or more students were administered seven or more tasks when started at Task 1; 97% or more students were administered nine or more tasks when started at Task 3. Similar patterns are seen in math, science, and social studies, with students starting at Task 3 showing large percentages going beyond the minimum number of tasks administered.

Generally, less than 5% of students across forms and subjects were not administered the minimum number of tasks required. The middle school social studies assessment presents

somewhat of an exception with 8% of the students starting at Task 3 having been administered fewer than the required nine tasks.

Students whose assessment started at Task 1 were administered between 8 and 9.3 tasks on average, their median number of administered tasks ranged between 7 and 9; students who started at Task 3 were administered between 10.9 and 11.9 tasks on average, with a median number of administered tasks between 11 and 13. These data indicate that, for both these groups of students, the tendency was to administer more than the minimum number of tasks needed. Students who started at Task 7 were administered 8.9 to 9 tasks on average, with a median of 9 tasks. In other words, students who started at Task 7 were generally administered all 9 tasks available at the high-complexity level.

These results show that a large majority of the students assessed during the 2011 spring SC-Alt administration were administered at least the minimum number of tasks, and in many instances the test administrators exposed the students to additional, more complex, and more difficult tasks beyond the minimal administration requirements.

Exhibit 8.8: Number of Tasks Administered, by Starting Task—ELA

		Number of Tasks Administered											
Starting Task		<6	6	7	8	9	10	11	12	>12	Total Students	Mean Number of Tasks	Median Number of Tasks
Elementary School													
1	N	2	.	195	33	41	76	4	4	24	379		
	%	0.5	0	51.5	8.7	10.8	20.1	1.1	1.1	6.3	100	8.47	7
3	N	3	2	1	4	71	23	10	12	180	306		
	%	1.0	0.7	0.3	1.3	23.2	7.5	3.3	3.9	58.8	100	11.52	13
7	N	4	.	3	1	788	.	.	.	0	796		
	%	0.5	0	0.4	0.1	99.0	0	0	0	0	100	8.97	9
Middle School													
1	N	3	.	164	21	20	1	8	10	55	282		
	%	1.1	0	58.2	7.5	7.1	0.4	2.8	3.6	19.5	100	8.99	7
3	N	1	2	3	.	43	29	14	7	145	244		
	%	0.4	0.8	1.2	0	17.6	11.9	5.7	2.9	59.4	100	11.63	13
7	N	3	2	1	3	787	.	.	.	0	796		
	%	0.4	0.3	0.1	0.4	98.9	0	0	0	0	100	8.96	9
High School													
1	N	0	.	47	4	2	4	4	2	12	75		
	%	0.0	0	62.7	5.3	2.7	5.3	5.3	2.7	16.0	100	8.84	7
3	N	0	.	.	1	19	7	2	5	36	70		
	%	0	0	0	1.4	27.1	10.0	2.9	7.1	51.4	100	11.41	13
7	N	1	.	.	.	209	.	.	.	0	210		
	%	0.5	0	0	0	99.5	0	0	0	0	100	8.97	9
Total													
	N	17	6	414	67	1980	140	42	40	452	3158		
	%	0.5	0.2	13.1	2.1	62.7	4.4	1.3	1.3	14.3	100	9.41	9

Exhibit 8.9: Number of Tasks Administered, by Starting Task—Mathematics

		Number of Tasks Administered											
Starting Task		<6	6	7	8	9	10	11	12	>12	Total Students	Mean Number of Tasks	Median Number of Tasks
Elementary School													
1	N	1	2	186	62	30	63	5	3	22	374		
	%	0.3	0.5	49.7	16.6	8.0	16.8	1.3	0.8	5.9	100	8.37	7
3	N	1	.	4	3	79	41	26	13	187	354		
	%	0.3	0	1.1	0.9	22.3	11.6	7.3	3.7	52.8	100	11.44	13
7	N	2	3	.	2	748	.	.	.	0	755		
	%	0.3	0.4	0	0.3	99.1	0	0	0	0	100	8.97	9
Middle School													
1	N	3	3	150	57	7	2	5	5	24	256		
	%	1.2	1.2	58.6	22.3	2.7	0.8	2.0	2.0	9.4	100	8.09	7
3	N	3	4	4	1	47	37	19	23	139	277		
	%	1.1	1.4	1.4	0.4	17.0	13.4	6.9	8.3	50.2	100	11.39	13
7	N	3	2	1	2	770	.	.	.	0	778		
	%	0.4	0.2	0.1	0.3	99.0	0	0	0	0	100	8.97	9
High School													
1	N	0	2	42	4	7	.	1	4	13	73		
	%	0	2.7	57.5	5.5	9.6	0	1.4	5.5	17.8	100	8.92	7
3	N	3	.	.	1	23	5	3	5	32	72		
	%	4.2	0	0	1.4	31.9	6.9	4.2	6.9	44.4	100	10.92	12
7	N	1	.	1	2	199	.	.	.	0	203		
	%	0.5	0	0.5	1.0	98.0	0	0	0	0	100	8.95	9
Total													
	N	17	16	388	134	1910	148	59	53	417	3142		
	%	0.5	0.5	12.4	4.3	60.8	4.7	1.9	1.7	13.3	100	9.36	9

Exhibit 8.10: Number of Tasks Administered, by Starting Task—Science/Biology

		Number of Tasks Administered											
Starting Task		<6	6	7	8	9	10	11	12	>12	Total Students	Mean Number of Tasks	Median Number of Tasks
Elementary School Science													
1	N	2	1	149	26	87	8	8	1	23	305		
	%	0.7	0.3	48.9	8.5	28.5	2.6	2.6	0.3	7.6	100	8.4	8
3	N	1	.	.	3	85	14	12	23	87	225		
	%	0.4	0	0	1.3	37.8	6.2	5.3	10.2	38.7	100	10.99	11
7	N	4	1	1	5	490	.	.	.	0	501		
	%	0.8	0.2	0.2	1	97.8	0	0	0	0	100	8.93	9
Middle School Science													
1	N	2	.	115	40	25	8	1	1	28	220		
	%	0.9	0	52.3	18.2	11.4	3.6	0.5	0.5	12.7	100	8.54	7
3	N	2	1	2	.	27	14	10	14	126	196		
	%	1.0	0.5	1.0	0	13.8	7.1	5.1	7.1	64.3	100	11.88	13
7	N	0	1	1	1	491	.	.	.	0	494		
	%	0	0.2	0.2	0.2	99.4	0	0	0	0	100	8.99	9
Total													
	N	11	4	268	75	1205	44	31	39	264	1941		
	%	0.6	0.2	13.8	3.9	62.1	2.3	1.6	2.0	13.6	100	9.35	9
High School Biology													
1	N	3	67	53	19	6	6	6	64		224		
	%	1.4	29.9	23.7	8.5	2.7	2.7	2.7	28.6		100	8.41	7
3	N	0	.	11	3	4	48	.	.		66		
	%	0	0	16.7	4.6	6.1	72.7	0	0		100	9.35	10
6	N	0	.	54		54		
	%	0	0	100	0	0	0	0	0		100	7	7
Total													
	N	3	67	118	22	10	54	6	64		344		
	%	0.9	19.5	34.3	6.4	2.9	15.7	1.7	18.6		100	8.37	7

Exhibit 8.11: Number of Tasks Administered, by Starting Task—Social Studies

		Number of Tasks Administered											
Starting Task		<6	6	7	8	9	10	11	12	>12	Total Students	Mean Number of Tasks	Median Number of Tasks
Elementary School													
1	N	0	2	115	23	15	17	29	1	31	233		
	%	0	0.96	49.4	9.9	6.4	7.3	12.5	0.4	13.3	100	8.98	7
3	N	1	2	.	.	79	7	4	6	87	186		
	%	0.5	1.1	0	0	42.5	3.8	2.2	3.2	46.8	100	10.99	11.5
7	N	4	.	1	1	597	.	.	.	0	603		
	%	0.7	0	0.2	0.2	99.0	0	0	0	0	100	8.97	9
Middle School													
1	N	1	1	95	21	21	4	2	.	10	155		
	%	0.7	0.7	61.3	13.6	13.6	2.6	1.3	0	6.5	100	8.03	7
3	N	1	.	9	3	37	14	19	.	77	160		
	%	0.6	0	5.6	1.9	23.1	8.8	11.9	0	48.1	100	11.09	11
7	N	0	4	3	1	594	.	.	.	0	602		
	%	0	0.7	0.5	0.2	98.7	0	0	0	0	100	8.97	9
Total													
	N	7	9	223	49	1343	42	54	7	205	1939		
	%	0.4	0.5	11.5	2.5	69.3	2.2	2.8	0.4	10.6	100	9.26	9

Number of Items Administered

Since test administrators were instructed to administer all of the items in a task and each task contained approximately five items, the number of items administered was roughly proportional to the number of tasks administered. Exhibits 8.12 through 8.15 show the mean, median, and 25th and 75th percentile for number of administered items, disaggregated by content area, form level, and task start point. Note that cases with non-standard starting tasks are not included in Exhibits 8.12–8.15.

The median number of items administered to students starting at Task 1 ranged between 29 and 45 across content areas and form levels; the median for Task 3 start points ranged between 52 and 79, and the median for Task 7 start points ranged between 44 and 64. Students beginning at Task 7 were administered fewer and a smaller range of items than students starting at Task 3 since these students demonstrated more predictable performance (according to the SPQ results) and the end of the minimally required task range coincided with the end of the test. Students whose test administration began at Task 3 tended to have more items administered to them. In these cases, the administration of a content area test was ended only when (1) the end of the test was reached or (2) the student could no longer respond successfully on a task (i.e., failed to obtain three or more points on the task).

Exhibit 8.12: Number of Items Administered, by Starting Task—ELA

Grade-Band	Start Task	N	Mean	P25	Median	P75
Elementary School	1	379	47.02	39	39	55
	3	306	64.38	50	73	73
	7	796	48.76	49	49	49
	TOTAL	1481	51.54	49	49	50
Middle School	1	282	48.26	38	38	60
	3	244	60.54	54	67	67
	7	796	44.76	45	45	45
	TOTAL	1322	48.42	45	45	47
High School	1	75	45.81	36	36	53
	3	70	58.71	45	67	68
	7	210	48.72	49	49	49
	TOTAL	355	50.08	49	49	49

Exhibit 8.13: Number of Items Administered, by Starting Task—Mathematics

Grade-Band	Start Task	N	Mean	P25	Median	P75
Elementary School	1	374	44.52	38	38	48
	3	354	58.13	47	65.5	66
	7	755	44.78	45	45	45
	TOTAL	1483	47.9	45	45	48
Middle School	1	256	38.38	33	33	37
	3	277	56.85	51	63	65
	7	778	45.78	46	46	46
	TOTAL	1311	46.67	46	46	46
High School	1	73	45.33	34	34	46
	3	72	60.13	50	66	72
	7	203	51.64	52	52	52
	TOTAL	348	52.07	50.5	52	52

Exhibit 8.14: Number of Items Administered, by Starting Task—Science/Biology

Grade-Band	Start Task	N	Mean	P25	Median	P75
Elementary School Science	1	305	42.28	35	40	45
	3	225	57.35	48	58	67
	7	501	45.63	46	46	46
	TOTAL	1031	47.2	45	46	48
Middle School Science	1	220	47.25	40	40	50
	3	196	62.67	59	68	68
	7	494	43.91	44	44	44
	TOTAL	910	48.76	44	44	49
High School Biology	1	224	45.67	33	38	64
	3	66	50.94	49	54	54
	6	54	38	38	38	38
	TOTAL	344	45.47	38	38	54

Exhibit 8.15: Number of Items Administered, by Starting Task—Social Studies

Grade-Band	Start Task	N	Mean	P25	Median	P75
Elementary School	1	233	44.13	32	32	55
	3	186	59.49	47	62.5	72
	7	603	51.76	52	52	52
	TOTAL	1022	51.43	51	52	52
Middle School	1	155	34.27	29	29	35
	3	160	52.53	41	52	63
	7	602	45.76	46	46	46
	TOTAL	917	45	45	46	46

Achievement Level of Students by Start Point

Within an SC-Alt form, two or more tasks (consisting of an average of five items each) are to be used to assess the same standards at different levels of student communication and content complexity, and they are to be placed on the test form in locations that ensure that there is adequate content coverage of the standards regardless of the student's starting point in the assessment. Although tasks are ordered on the form based on student communication levels and average content complexity, items of both lower and higher complexity may appear in each task. This configuration presents items and tasks across the entire assessment providing students with opportunities to demonstrate proficiency. Each student's proficiency and resulting achievement level are determined by the student's performance on the specific group of items the student was administered. The calculation of student proficiency scores is described in Chapter 5. The distribution of achievement levels for students according to start task, form level, and content area is presented in Exhibit 8.16.

The table entries demonstrate interesting operational aspects of the leveled structure of the SC-Alt. Across content areas, students beginning the assessment at Task 1 are categorized as Proficient (achievement Levels 3 and 4) at rates between 5% and 47%, with the lowest percentage in middle school (5%–18%), and more varied in elementary school (8%–40%) and high school (11%–47%). For students starting at Task 3, 33% to 92% of students across content areas are categorized as Proficient; as with students beginning with Task 1, the lowest percentage of Proficient students starting at Task 3 is demonstrated in middle school (33%–68%), while a wide range is seen in both elementary school (39%–92%) and high school (39%–86%). Finally, 62% to 100% of students starting at Task 6 or 7 are Proficient.

Summary

The purpose of the start and stop point analyses was to document the number of tasks and items students complete during the assessment and the performance levels that groups of students attain who enter the assessment at different start points. The results of these analyses demonstrate the effectiveness of the use of the SPQ and the test administration start/stop rules that are based on the student's performance during the assessment. Except in a few instances, all students were administered adequate numbers of tasks and items to assess the intended content.

The findings indicate that SPQ start/stop rules were being followed for almost all tested students. A considerable proportion of tested students continued testing beyond the minimum required number of tasks to be administered. As a consequence, in each starting task group, there were students who tested in the proficient range (i.e., at achievement Levels 3 or 4). Finally, students assigned to higher starting tasks showed a greater likelihood of testing in the proficient range. These results demonstrate that the tailored assessment design of the SC-Alt operates as expected.

Exhibit 8.16: Achievement Level by Task Start Point, Form Level, and Content Area

Subject	Ach. Level	Elementary School (ES)								Middle School (MS)								High School (HS)							
		Starting Task								Starting Task								Starting Task							
		1		3		7		TOTAL		1		3		7		TOTAL		1		3		6/7 ¹⁰		TOTAL	
		N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
ELA	Level 1	75	19.8	.	0.0	.	0.0	75	5.1	73	25.9	1	0.4	.	0.0	74	5.6	24	32.0	.	0.0	.	0.0	24	6.8
	Level 2	213	56.2	73	23.9	11	1.4	297	20.1	159	56.4	78	32.0	21	2.6	258	19.5	38	50.7	32	45.7	11	5.2	81	22.8
	Level 3	77	20.3	131	42.8	108	13.6	316	21.3	40	14.2	96	39.3	94	11.8	230	17.4	6	8.0	23	32.9	26	12.4	55	15.5
	Level 4	14	3.7	102	33.3	677	85.1	793	53.5	10	3.6	69	28.3	681	85.6	760	57.5	7	9.3	15	21.4	173	82.4	195	54.9
	Proficient	91	24.0	233	76.1	785	98.6	1109	74.9	50	17.7	165	67.6	775	97.4	990	74.9	13	17.3	38	54.3	199	94.8	250	70.4
Math	Level 1	114	30.5	1	0.3	.	0.0	115	7.8	89	34.8	2	0.7	1	0.1	92	7.0	27	37.0	2	2.8	.	0.0	29	8.3
	Level 2	208	55.6	120	33.9	13	1.7	341	23.0	148	57.8	131	47.3	27	3.5	306	23.3	38	52.1	42	58.3	18	8.9	98	28.2
	Level 3	50	13.4	200	56.5	271	35.9	521	35.1	17	6.6	123	44.4	305	39.2	445	33.9	6	8.2	22	30.6	108	53.2	136	39.1
	Level 4	2	0.5	33	9.3	471	62.4	506	34.1	2	0.8	21	7.6	445	57.2	468	35.7	2	2.7	6	8.3	77	37.9	85	24.4
	Proficient	52	13.9	233	65.8	742	98.3	1027	69.3	19	7.4	144	52.0	750	96.4	913	69.6	8	11.0	28	38.9	185	91.1	221	63.5
Science/ Biology	Level 1	86	28.2	.	0.0	.	0.0	86	8.3	110	50.0	6	3.1	.	0.0	116	12.8	36	16.1	.	0.0	.	0.0	36	10.5
	Level 2	98	32.1	18	8.0	2	0.4	118	11.5	77	35.0	74	37.8	31	6.3	182	20.0	83	37.1	9	13.6	3	5.6	95	27.6
	Level 3	78	25.6	76	33.8	49	9.8	203	19.7	21	9.6	72	36.7	78	15.8	171	18.8	46	20.5	12	18.2	9	16.7	67	19.5
	Level 4	43	14.1	131	58.2	450	89.8	624	60.5	12	5.5	44	22.5	385	77.9	441	48.5	59	26.3	45	68.2	42	77.8	146	42.4
	Proficient	121	39.7	207	92.0	499	99.6	827	80.2	33	15.0	116	59.2	463	93.7	612	67.3	105	46.9	57	86.4	51	94.4	213	61.9
Social Studies	Level 1	87	37.3	1	0.5	.	0.0	88	8.6	95	61.3	4	2.5	.	0.0	99	10.8								
	Level 2	128	54.9	113	60.8	39	6.5	280	27.4	52	33.6	104	65.0	74	12.3	230	25.1								
	Level 3	13	5.6	60	32.3	272	45.1	345	33.8	6	3.9	44	27.5	272	45.2	322	35.1								
	Level 4	5	2.2	12	6.5	292	48.4	309	30.2	2	1.3	8	5.0	256	42.5	266	29.0								
	Proficient	18	7.7	72	38.7	564	93.5	654	64.0	8	5.2	52	32.5	528	87.7	588	64.1								

¹⁰ Starting Task 6 for high school biology, Task 7 for ELA and mathematics.

References

- Agresti, A. (1990). *Categorical data analysis*. New York: Wiley.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- American Institutes for Research. (2007). *South Carolina Alternate Assessment spring 2007 standard setting technical report*. Washington, DC: American Institutes for Research.
- American Institutes for Research. (2008). *South Carolina Alternate Assessment (SC-Alt): Technical report for English language arts and mathematics field test administration, spring 2006*. Washington, DC: American Institutes for Research.
- American Institutes for Research and South Carolina Department of Education. (2008). *South Carolina's Alternate Assessment, SC-Alt spring 2007 operational administration*. Washington, DC: American Institutes for Research.
- American Institutes for Research and South Carolina Department of Education. (2009). *South Carolina's Alternate Assessment, SC-Alt spring 2008 operational administration*. Washington, DC: American Institutes for Research.
- American Institutes for Research and South Carolina Department of Education. (2010a). *South Carolina's Alternate Assessment, SC-Alt spring 2010 operational administration*. Washington, DC: American Institutes for Research.
- American Institutes for Research and South Carolina Department of Education. (2010b). *South Carolina Alternate Assessment 2010 standard setting: Setting standards in high school biology technical report*. Washington, DC: American Institutes for Research.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81–105.
- Center for Research on Education. (2009a). *Alternate assessment alignment study of replacement items report to the South Carolina State Department of Education*. Columbia, SC: South Carolina State Department of Education.
- Center for Research on Education. (2009b). *High school alternate assessment alignment report to the South Carolina State Department of Education*. Columbia, SC: South Carolina State Department of Education.
- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage Publications.
- Education Oversight Committee. (2008a). *Review of the SC-Alternate English language arts and mathematics assessments report to the South Carolina State Department of Education*. Columbia, SC: South Carolina State Department of Education.

- Education Oversight Committee. (2008b). Review of the SC-Alternate science assessment report to the South Carolina State Department of Education. Columbia, SC: South Carolina State Department of Education.
- Ferrara, S., Perie, M., & Johnson, E. (2008). Matching the judgmental task with standard setting panelist expertise: The Item-Descriptor (ID) Matching procedure. *Journal of Applied Testing Technology*, 9, 1–22.
- Fiske, D. W. (1995). Reprise, new themes, and steps forward. In P. E. ShROUT & S. T. Fiske (Eds.), *Personality research, methods, and theory: A festschrift honoring Donald W. Fiske*. Hillsdale, NJ: Lawrence Erlbaum.
- Flowers, C., Browder, D., Wakeman, S., & Karvonen, M. (2006a). *Alternate assessment alignment pilot study report to the South Carolina State Department of Education*. Columbia, SC: South Carolina State Department of Education.
- Flowers, C., Browder, D., Wakeman, S., & Karvonen, M. (2006b). *Science alternate assessment alignment pilot study report to the South Carolina State Department of Education*. Columbia, SC: South Carolina State Department of Education.
- Flowers, C., Browder, D., Wakeman, S., & Karvonen, M. (2007). Science alternate assessment alignment pilot study report to the South Carolina State Department of Education, Addendum, December 21, 2007. Columbia, SC: South Carolina State Department of Education.
- Huynh, H. (2003, August). *Technical memorandum for computing standard error in Bookmark standard setting* [The South Carolina PACT 2003 Standard Setting Support Project]. Columbia, SC: University of South Carolina.
- Kleinert, H. L., & Kearns, J. F. (2001). *Alternate assessment: Measuring outcomes and supports for students with disabilities*. Baltimore, MD: Brookes.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149–174.
- National Alternate Assessment Center. (2005). *Designing from the ground floor*. Conference Pre-Session, Council of Chief State School Officers' Large-Scale Assessment Conference. San Antonio, TX: June 17–18, 2005.
- Rudner, L.M., & Schafer, W.D. (2001), Reliability (ERIC Digest). College Park, MD: ERIC Clearinghouse on Assessment and Evaluation (ERIC Document Reproduction Service No. ED458213). Retrieved May 17, 2009, from www.ericdigests.org/2002-2/reliability.htm.
- Sireci, S. G., Thissen, D., & Wainer, H. (1991). On the reliability of testlet-based tests. *Journal of Educational Measurement*, 28(3), 234–247.
- Spitzer, R., Cohen, J., Fleiss, J., & Endicott, J. (1967). Quantification of agreement in psychiatry diagnosis: A new approach. *Archives of General Psychiatry*, 17, 83–87.

Appendices

Appendix A: Assignment of Tasks to Grade-Band Forms for the Spring 2011 Administration

All tasks in each SC-Alt grade-band assessment align with the extended standards in that grade-band. Because adjacent grade-band score scales are linked psychometrically, some tasks in each grade-band assessment align with extended standards in both adjacent grade-bands. In turn, the extended standards link back to separate grade-level performance standards for the appropriate grades.

All items in linking tasks are developed to be appropriate for students in both adjacent grade-bands. In some cases (e.g., some tasks in ELA), the extended standards to which linking tasks align are equivalent for two adjacent grade-bands. However, the grade-level performance standards to which the extended standards are linked *do* differ across the adjacent grade-bands. In all content areas and for all grade-bands, Descriptions of Achievement Levels (DALs) are specific to each grade-band and differ across grade-bands.

2011 Operational Test Designs to Support Psychometric Linking of Grade-Band Score Scales

To provide data to link all grade-band assessments onto a vertical scale, linking tasks were repeated in adjacent grade-band assessments. For example, three of the tasks that appeared in the ELA grade-band 3–5 assessment also appeared in the ELA grade-band 6–8 assessment. Those three linking tasks and the nine unique tasks made up the 12 tasks in the ELA grade-band 3–5 assessment. The ELA grade-band 6–8 assessment included the three linking tasks from the grade-band 3–5 assessment, five linking tasks that also appeared in the grade-band 9–12 assessment, and four unique tasks. This “linking upward” design ensures that students were assessed on tasks aligned with their current grade placement or previous grades.

English Language Arts Assessment

The ELA assessment covered extended standards in reading, writing, and communication. The 2011 operational test forms included 12 operational tasks for each of three grade-band assessments. Eight of these tasks were used to link between adjacent grade-band assessments, so there was a total of 28 tasks. The design for the ELA assessment for spring 2011 appears in Exhibit A-1.

Exhibit A-1: Numbers of Tasks in Each Operational Grade-Band Assessment, ELA

Grade-Band	Unique Tasks	Linking Tasks		Total for Operational Test
	Tasks in Each Grade-Band Assessment Test Booklet			
3–5	9	3	—	12
6–8	4		5	12
10	7	—		12
	Tasks to Be Included			
All grades	20	8		36

Mathematics Assessment

The mathematics assessment covered the mathematics standards. The 2011 operational test forms included 12 tasks for each of three grade-band assessments, with 17 unique and eight linking tasks, for a total of 25 tasks. The design of the mathematics assessment for spring 2011 appears in Exhibit A-2.

Exhibit A-2: Numbers of Tasks in Each Operational Grade-Band Assessment, Mathematics

Grade-Band	Unique Tasks	Linking Tasks		Total for Operational Test	
	Tasks in Each Grade-Band Assessment Test Booklet				
3–5	5	4	—	3	12
6–8	4		1		12
10	8				—
	Tasks to Be Included				
All grades	17	8		36	

Science/Biology Assessment

The science assessment covered the science extended standards. The 2011 operational test forms included 12 tasks for each of three grade-band assessments for a total of 36 tasks. The test assesses physical science content in the elementary school and middle school grade-bands and biology in the high school grade-band. The physical science items on the elementary and middle school forms are calibrated to the same operational item bank and are vertically linking. However, there are no linking tasks on the 2011 science forms.

Social Studies Assessment

The social studies assessment covered the social studies standards. The 2011 operational test forms includes 12 tasks for each of two grade-band assessments, with 16 unique and four linking tasks, resulting in a total of 20 tasks. The design for the social studies assessment for spring 2011 appears in Exhibit A-3.

Exhibit A-3: Numbers of Tasks in Each Operational Grade-Band Assessment, Social Studies

Grade-Band	Unique Tasks	Linking Tasks	Total for Operational Test
	Tasks in Each Grade-Band Assessment Test Booklet		
3–5	8	4	12
6–8	8		12
	Tasks to Be Included		
All grades	16	4	24

Appendix B: Starting and Stopping Rules for Using the Student Placement Questionnaire

Directions for Determining the Starting and Concluding Tasks and Use of the Student Placement Questionnaire, Spring 2011

These directions guide you through the following steps:

- Completing the Student Placement Questionnaire (SPQ),
- Identifying the starting task in each content area,
- Adjusting the starting task, if that becomes necessary,
- Determining when to conclude the administration

Completing the Student Placement Questionnaire

The SPQ is designed to identify the most appropriate starting task for each of your students in each content area of SC-Alt. You will use the SPQ to identify the most appropriate starting task for each student in the SC-Alt assessments in English language arts, mathematics, science, and social studies. Answer each SPQ item as accurately as you can based on your experience in the classroom with this student.

The SPQs are located in the Student Answer Folder along with the areas for recording the student's scores on each SC-Alt task. An example of a completed English Language Arts SPQ is included at the end of these instructions.

Identifying the Starting Task for a Student in Each Content Area

1. Bubble in your responses to the SPQ questions.

After you respond to all items in the SPQ, identify the most appropriate starting task for this student following the steps on the SPQ. These are the steps:

2. Count the number of bubbles you marked in each of the first three columns, and write the totals in the blocks under each column.
3. In section 3 at the bottom of the page:
 - a. Write the column totals in the appropriate blocks.
 - b. Multiply each total by the specified multiplier, and write the resulting totals in the blocks to the right.
 - c. Sum the three totals to obtain the total SPQ score. Write the SPQ score into the blocks and bubble in the SPQ score.
 - Please check your work and complete the bubble grids for the total SPQ score.
4. Find the total SPQ score in section 4 to determine the starting task for this student.

Administering the Starting Task and Completing the Administration

After you identify the starting task for this student using the SPQ, follow these directions to administer the starting task and complete the administration.

The SPQ provides the initial starting point for a student's administration. Each student must be administered a minimum of six operational tasks (including the starting task) if the student is started at Task 1 or a minimum of nine tasks if the student is started at Task 3 or Task 7 (or Task 6 for high school biology). The minimum number of tasks and specific tasks that must be administered to each student for each starting level are specified in the table below.

Exhibit B-1: Minimum Task Ranges to Be Administered

ELA, Mathematics, Science, and Social Studies	
Starting task	Administer all items in at least these tasks
Task 1	1–7
Task 3	3–11
Task 7	7–15

Please note: The Biology assessment contains 12 tasks. The minimum number of Biology tasks and specific tasks that must be administered to each student for each starting level are specified in the table below.

Biology	
Starting task	Administer all items in at least these tasks
Task 1	1–6
Task 3	3–9
Task 6	6–12

It may be necessary to adjust the starting task based on the student's level of success on the first task. Also, the administration should be continued beyond the minimum number of tasks when the student is responding successfully.

When the Student Does Not Respond Successfully on the First Task

“Responding successfully” means getting at least three total points on a task. Each task has at least four items. Responding successfully would mean that a student received at least three total points for all the items combined. For example, a student may respond successfully by receiving three points on one item, or two points on one item and one point on another item, or one point each on three different items. When a student does not receive three or more total points on a task, the student has not responded successfully on the task.

When a student is started at Task 3 or at Task 6/7 and does not respond successfully on the first task, the starting task was too difficult, and the teacher must restart the student at the next lower starting point. For example:

- If the student starts at Task 3 but cannot respond successfully on Task 3, restart the student at Task 1.
- If the student starts at Task 6/7 but cannot respond successfully on Task 6/7, restart the student at Task 3.

When a student is started at Task 1, no downward adjustment is possible, and the administration must progress through at least five tasks.

When to Conclude the Administration

If the student responds successfully on the last required task as specified in the table above, continue with the administration by administering the next task and subsequent tasks until the student no longer responds successfully on a task. By continuing the administration of subsequent tasks when the student is “responding successfully,” you will provide the maximum opportunity for the student to demonstrate his or her knowledge and skills.

If the student does not respond successfully on the last required task or if at any point the student does not respond successfully on additional tasks (i.e., obtain three or more points on the task), you may conclude the administration.

By concluding the administration when the student is no longer successful after you have administered the required tasks, the student’s test administration is not prolonged unnecessarily, and possible negative effects on the student are avoided.

Examples:

- Student A was started at ELA Task 1 and administered Tasks 1–7. The student responded successfully on Task 7 and therefore was administered Task 8. The student responded successfully on Task 8 and was administered Task 9. The student did not respond successfully on Task 9, and the administration was concluded after Task 9.
- Student B was started at Task 3 and was administered Tasks 3–11. The student did not respond successfully on Task 11, and the administration was concluded after Task 11.

SC - ALT STUDENT PLACEMENT QUESTIONNAIRE - ELA (ENGLISH LANGUAGE ARTS)

Follow steps 1-4 to complete the SPQ and identify the starting task.

(1) Please darken the bubble (●) that corresponds to the most appropriate response for this student. Mark only one response for each item. Please mark a response for all items below. Use a No. 2 pencil only.

	No, she/he cannot do this
With physical prompting/hand-over-hand	
With verbal/gestural prompting	
Independently	

In reading, can this student:

- | | | | | |
|---|-----------------------|----------------------------------|----------------------------------|----------------------------------|
| 1. Attend to text read aloud? | <input type="radio"/> | <input checked="" type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 2. Recall details of text read aloud? | <input type="radio"/> | <input checked="" type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 3. Recognize some high-frequency written words? | <input type="radio"/> | <input type="radio"/> | <input checked="" type="radio"/> | <input type="radio"/> |
| 4. Draw conclusions or make inferences about texts? | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input checked="" type="radio"/> |

In writing, can this student:

- | | | | | |
|---|-----------------------|----------------------------------|----------------------------------|----------------------------------|
| 5. Write his or her name using a pencil, name stamp, letter titles, or other means? | <input type="radio"/> | <input checked="" type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 6. Use objects, pictures, and/or picture symbols to write in any format? | <input type="radio"/> | <input type="radio"/> | <input checked="" type="radio"/> | <input type="radio"/> |
| 7. Copy, trace, or print letters? | <input type="radio"/> | <input checked="" type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 8. Use oral language and/or letters and words to write? | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input checked="" type="radio"/> |

In communicating, can this student:

- | | | | | |
|--|----------------------------------|----------------------------------|-----------------------|-----------------------|
| 9. Listen (i.e., demonstrate receptive behavior) and respond? | <input checked="" type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 10. Participate in conversations by responding appropriately? | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 11. Use language to express a preference, opinion, or viewpoint? | <input type="radio"/> | <input checked="" type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 12. Recognize and understand the meaning of environmental signs (e.g., street signs, store signs, school signs)? | <input type="radio"/> | <input checked="" type="radio"/> | <input type="radio"/> | <input type="radio"/> |

(2) Write in the total number of bubbles you marked in each column

1 7 2

Col.1 Col.2 Col.3

(3) Calculate the SPQ total score

- (a) write the column totals from (2) in (a) below
 (b) multiply and write the results in (b) below
 (c) sum the results from (b) and write the sum in (c)

	(a)		(b)
Column 1 Total	1	x 3 =	3
			+
Column 2 Total	7	x 2 =	14
			+
Column 3 Total	2	x 1 =	2
(c) Total SPQ Score = 19			

(4) Identify the starting task for this student using the SPQ total score from step (3).

If the total score is in this range	Start at this task	Administer all items in <u>at least</u> these tasks
0-11	Task 1	1-7
12-22	Task 3	3-11
23-36	Task 7	7-15

Appendix C: Scoring Audits and Analysis of Video Rater and Second Rater Data from the Spring 2011 Operational Administration

A videotaping and a second rater study were conducted to audit scoring accuracy and classification consistency for the spring 2011 administrations of the SC-Alt in ELA and Biology (i.e., high school science). **Scoring accuracy** refers to the degree to which teachers follow scaffolding and scoring directions correctly and assign correct scores to student responses. **Classification consistency** refers to the degree to which students are assigned to the same performance levels based on item scores by test administrators and second raters or video raters. This appendix describes the sampling procedures, the identified sample of students, and the results for the attained sample of completed videotaped and second rater administrations.

Sampling Procedures

The sampling procedure was designed to include administrations from every school district and to be broadly representative of the range of student and test administrations. A sample of students was identified for videotaping so that (a) all districts implementing the SC-Alt would be required to videotape at least one student administration and (b) the total number of taped administrations per district would be based on the number of teachers involved in the assessment for each district. The sampling was by teacher and student within districts. One-third of the teachers within each district were randomly sampled to videotape the administration of one student. A subset of elementary students from the videotape sample was selected for a second rater pilot. These students had their item responses scored simultaneously by a second rater who was present during the test administration. This pilot sample was not videotaped. The number of teachers (and students) to be selected from each district is shown in Exhibit C-1.

Exhibit C-1: District-Based Sampling Targets for Video Rater Study

Total Number of Teachers per District	Number Required to Videotape
1–5	1
6–8	2
9–11	3
12–14	4
15–17	5
18–20	6
21–23	7
24–26	8
27–29	9
30–32	10
33–35	11
36–38	12
39–41	13
42–44	14
45–47	15
48–50	16
51–53	17
54–56	18
57–59	19
60–62	20

Based on this sampling plan and the numbers of pre-identified students coded for each district for the 2011 administration, the frequency distribution of test administrations sampled per district was as follows:

ELA: Second Rater (Elementary School only)

- 1 test administration – 12 districts
- 2–5 test administrations – 12 districts

ELA: Video Rater (All Grade-Bands)

- 1 test administration – 35 districts
- 2–5 test administrations – 26 districts
- 6–10 test administrations – 8 districts
- 11–15 test administrations – 2 districts
- 16–20 test administrations – 1 district

Biology: Video Rater (High School only)

- 1 test administration – 20 districts
- 2–5 test administrations – 12 districts
- 6–10 test administrations – 3 districts

The sampling of students and teachers was conducted from the January 2011 precode file, which was the pre-identification file for the spring 2011 SC-Alt administration. The sampling

was conducted by SCDE, and the students identified for videotaping were flagged on the precode file sent to AIR for the production of materials and district notification.

Videotaping Procedures

The district test coordinators for alternate assessment were provided rosters of the students identified for videotaping. The district materials included a packet of information for each teacher that included the following:

- A videotaping student roster identifying the student
- Information on the purpose of the videotaping and instructions for conducting the videotaped administrations
- A videotaping student information form
- Barcode labels for positive identification and linking of the videotapes and the student information to the sc-alt assessment data file
- Directions for the packaging and return of materials

Communications to teachers and district test coordinators emphasized the importance of completing the videotaped administrations, provided contact information for questions or concerns, and asked for notification of SCDE if there were any problems in completing a videotaped administration for a particular student. Districts notified SCDE about a small number of students who either could not be assessed (e.g., because the students had moved, the parents did not consent to videotaping, or the students were not going to be assessed with the SC-Alt) or for whom the videotaping was inappropriate or extremely difficult to implement (e.g., medical homebound students or students whose behavior would become disruptive as a result of taping). These students were deleted from the videotaping sample.

Approximately one-half of the students who were deleted from the videotaping sample list by SCDE were replaced by another student with the same teacher or in a few cases by identifying a different teacher and student. The replacement students were selected to match the grade-band form and disability of the original students as closely as possible. As a result of notifications by districts, the SCDE deleted 43 students from the original sample and instructed districts to videotape 18 replacement students.

Analysis of Video Rater (VR) Data

The total number of students identified for videotaping and second rater assessment after SCDE adjustments (resulting from deletions and replacements) was 281. Videotaping materials were received for 218 ELA and 54 biology assessments. Of the ELA video materials, scorable VR assessments were obtained for 175 cases. Similarly, 41 of the 54 biology video data yielded scorable assessments.

Analysis of Second Rater (SR) Data

During the spring 2011 administration of the SC-Alt, 53 cases of second rater data were collected from elementary school ELA administrations. These administrations involved a second-rater observer independently scoring the administration along with the test administrator (TA). The scoring data from the second rater was recorded on a separate answer folder and the data was compared to the official TA scoring data to complete the same scoring consistency and classification consistency analyses that were conducted for the videotaping data. By obtaining

samples of both second rater procedure and videotaping procedure data from elementary ELA administrations, the effectiveness of the two procedures could be compared.

Five of the 53 returned SR booklets lacked unique student identifying information, preventing their being matched to TA booklets. After setting these five SR booklets aside, a total of 48 SR booklets remained for analysis.

Exhibit C-2 displays the student demographics of the video and second rater studies broken down by content area, grade-band, and data source (VR vs. SR).

Exhibit C-2: Demographic Frequencies for the Video Rater and Second Rater Data Samples—by Test Form

	ELA								Biology	
	Elementary School				Middle School		High School		High School	
	Second Rater		Video Rater		Video Rater		Video Rater		Video Rater	
	N	%	N	%	N	%	N	%	N	%
STUDENT'S ETHNICITY										
African American	23	47.9	39	52	47	53.4	34	61.8	34	63.0
American Indian	.	0	.	0	.	0	1	1.8	.	0
Asian	.	0	3	4	1	1.14	.	0	.	0
Hawaiian/Pacific Islander	.	0	.	0	.	0	.	0	.	0
Hispanic	4	8.3	.	0	2	2.3	.	0	.	0
Other	1	2.1	1	1.3	2	2.3	.	0	.	0
White	20	41.7	32	42.7	36	40.9	20	36.4	20	37.0
STUDENT'S GENDER										
Female	11	22.9	25	33.3	26	29.6	16	29.1	16	29.6
Male	37	77.1	50	66.7	62	70.5	39	70.9	38	70.4
ESL (LANGUAGE)										
Advanced	.	0	.	0	.	0	.	0	.	0
Advanced Waiver	.	0	.	0	.	0	.	0	.	0
Beginner	.	0	.	0	.	0	.	0	.	0
Beginner Waiver	.	0	.	0	.	0	.	0	.	0
English Speaker I	.	0	.	0	1	1.1	.	0	.	0
English Speaker II	45	93.8	73	97.3	84	95.5	55	100	54	100
Full English Proficient	.	0	.	0	.	0	.	0	.	0
Intermediate	.	0	.	0	.	0	.	0	.	0
Intermediate Waiver	.	0	.	0	.	0	.	0	.	0
Pre-functional	3	6.3	2	2.7	3	3.4	.	0	.	0
Pre-functional Waiver	.	0	.	0	.	0	.	0	.	0
Title III First Year Exited	.	0	.	0	.	0	.	0	.	0
Title III Second+ Year Exited	.	0	.	0	.	0	.	0	.	0
Unknown	.	0	.	0	.	0	.	0	.	0

	ELA								Biology	
	Elementary School				Middle School		High School		High School	
	Second Rater		Video Rater		Video Rater		Video Rater		Video Rater	
	N	%	N	%	N	%	N	%	N	%
ELIGIBLE FOR FREE OR REDUCED-PRICE LUNCH										
Full-Pay Meals	16	33.3	25	33.3	21	23.9	21	38.2	20	37.0
Free Meals	30	62.5	46	61.3	62	70.5	32	58.2	32	59.3
Reduced Meals	2	4.2	4	5.3	5	5.7	2	3.6	2	3.7
EFA GRADE (REPORTED GRADE FOR FUNDING)										
1	1	2.1	.	0	.	0	.	0		
2	2	4.2	7	9.3	.	0	.	0		
3	14	29.2	26	34.7	.	0	.	0		
4	21	43.8	27	36.0	2	2.3	.	0		
5	10	20.8	15	20.0	15	17.1	.	0		
6	.	0	.	0	29	33.0	.	0		
7	.	0	.	0	26	30.0	1	1.8	1	1.9
8	.	0	.	0	13	14.8	3	5.5	3	5.6
9	.	0	.	0	3	3.4	22	40.0	22	40.7
10	.	0	.	0	.	0	28	50.9	27	50.0
11	.	0	.	0	.	0	1	1.8	1	1.9
COMPLETION STATUS: Student satisfied attemptedness rule										
ELA	48	100	70	93.3	61	70.1	44	80.0		
Biology									41	75.9
COMPLETION STATUS: Student did not satisfy attemptedness rule										
ELA	.	0	1	1.3	4	4.6	.	0		
Biology									1	1.9
COMPLETION STATUS: Not scorable due to video issues										
ELA	.	0	4	5.3	22	25.3	11	20.0		
Biology									12	22.2
Student Residence										
Migrant Status	.	0	.	0	.	0	.	0	.	0
Home Schooled	.	0	.	0	.	0	.	0	.	0
Medical Homebound	.	0	.	0	.	0	.	0	.	0
IEP Disability Codes (Multiple Codes per Student)										
Severely Mentally Disabled	6	12.5	5	6.7	8	9.1	3	5.5	3	5.6
Moderately Mentally Disabled	5	10.4	15	20	29	33.0	24	43.6	24	44.4
Mildly Mentally Disabled	12	25	16	21.3	23	26.1	17	30.9	17	31.5
Autism	16	33.3	18	24	21	23.9	8	14.6	8	14.8

	ELA								Biology	
	Elementary School				Middle School		High School		High School	
	Second Rater		Video Rater		Video Rater		Video Rater		Video Rater	
	N	%	N	%	N	%	N	%	N	%
Deaf/Blindness	.	0	.	0	.	0	.	0	.	0
Emotional Disability	.	0	1	1.3	.	0	.	0	.	0
Hearing Impaired	1	2.1	1	1.3	1	1.1	2	3.6	2	3.7
Learning Disability	2	4.2	4	5.3	2	2.3	2	3.6	2	3.7
Multiple-Disability	.	0	.	0	.	0	.	0	.	0
Other Health Impaired	1	2.1	7	9.3	5	5.7	4	7.3	3	5.6
Orthopedically Impaired	2	4.17	3	4	6	6.82	1	1.82	1	1.85
Speech or Language Impaired	36	75	59	78.7	38	43.2	20	36.4	19	35.2
Traumatic Brain Injury	1	2.1	.	0	.	0	.	0	.	0
Visually Impaired	2	4.2	4	5.3	2	2.3	2	3.6	2	3.7
TOTAL	48	100	75	100	88	100	55	100	54	100

When the attained video and second rater samples are compared to the assessed population (see Exhibit 3.3), the following statements can be made:

By Form

- The attained sample approximates the expected number of students for each form: Elementary students (SR and VR samples combined) make up 46% of the sample, middle school students 33%, and high school students make up 19% of the sample.

By IEP Disability Code

The first four rows show the primary disabilities of severe, moderate, and mild mental disability and autism. If any of the mental disabilities were coded together with autism, then only the mental disability was reported. Subsequent rows show additional disabilities coded by the test administrators. Since multiple disability codes per student are permitted, the percentages do not add up to 100.

- Severe Mental Disability was sampled similarly to the expectation across forms (elementary school: 6.7%, middle school: 9.1%, and high school: 5.5%).
- Moderate Mental Disability was sampled at a similar rate (20%, 33%, and 44%) to the expectation.
- Mild Mental Disability was sampled at a similar rate (21%, 26%, and 31%) to the expectation.
- Autism was sampled at a similar rate (24%, 24%, and 15%) to the expectation.
- The total percentages of students in primary disability categories other than Severe, Moderate, and Mild Disability and Autism were represented at lower rates in elementary and middle school and a similar rate in high school compared with the identified sample (28%, 8%, and 5%).

By Other Demographic Variables

For other demographic variables, the proportions in the attained VR and SR samples generally appear to correspond to those seen in the total assessed population when data were available.

- In the sample, African American (48%–63%),¹¹ American Indian (0%–2%), Asian (0%–4%), Hispanic (0%–8%), White (37%–43%), and Other (0%–2%) ethnicities were reported, representing the majority of ethnicities in the total assessed population. These percentages evidence some variability around the corresponding population values as a result of the small sample sizes.
- Gender is distributed as approximately two to three males for each female.
- “English Speaker II” (94% to 100%) in the sample reflects the percentage of students in the assessed population.
- Between 58% and 71% of students in the sample were eligible for Free Lunch, approximately the same as in the total population. A small group of students in the sample was eligible for Reduced Lunch (4% to 5%), which is similar to the assessed population (6% to 7%).
- None of the students in the attained VR and SR samples were home-schooled, migrant, or medically homebound.

The attained VR and SR sample (Exhibit C-2) appears to reasonably represent the full population (Exhibit 3.3). The demographic variables of interest are present in the attained sample data within acceptable ranges of the assessed population.

Item Agreement Analysis

Within each grade-band, the absolute difference between test administrator (TA) scores and AIR video rater (VR) scores for each item, or second rater (SR) item scores, was computed. Scores that did not differ between TA and VR/SR are noted as “equal”; scores differing by ± 1 score point were noted as “adjacent”. Scores differing by more than ± 1 point were flagged as “discrepant.” The agreement data are summarized by content area and grade-band in Exhibit C-3, where values indicate the average percentage of items falling within each agreement category for which there were valid matched responses across TAs and VRs/SRs.

For the elementary school ELA form, the VR audit showed 96% of items scored as “equal” between the TA and VR; “adjacent” ratings were the next most prevalent outcome (at 4%); and “discrepant” ratings were the least prevalent result for all content area areas (less than 1%). The SR audit provided similar results, with 97% of the item scores equal, 3% adjacent, and 0.5% discrepant.

On the middle school form, the VR study showed a pattern similar to the elementary form: 95% of the ELA item scores matched as “equal,” 4% as “adjacent,” and 1% as “discrepant.”

On the high school form, the VR study yielded the same pattern for both ELA and Biology: “equal” ratings again account for the largest proportion of cases (94% to 96%), “adjacent” is the next most prevalent (4% to 5%), and finally “discrepant” (0% to 1%).

¹¹ The percentage range is reported across all five subsamples—elementary school ELA SR, elementary, middle, and high school ELA VR, and high school biology VR.

In comparison, regarding the second rater and video rater studies of the elementary school ELA assessment, the agreement statistics from the two samples are nearly the same. In other words, the second rater and video studies yield the same agreement statistics.

Exhibit C-3: Average Item Agreement Statistics by Method, Grade-Band and Subject

Subject	Agreement	Second Rater		Video Rater					
		Elementary School		Elementary School		Middle School		High School	
		Response Count	%	Response Count	%	Response Count	%	Response Count	%
ELA	Equal	2085	96.9	2539	96.0	2429	94.9	1679	95.9
	Adjacent	56	2.6	126	3.6	88	4.2	63	3.5
	Discrepant	10	0.5	26	0.4	15	0.9	7	0.6
Biology	Equal							1627	94.3
	Adjacent							93	5.4
	Discrepant							6	0.3

Classification Consistency Analysis (as distinct from scoring consistency as discussed in the previous section)

The reported performance levels for each student are derived from a scale score to performance level conversion process. Scale scores are produced based on conversions from the raw scores assigned by the TA. From these scale scores, students were assigned to one of four performance levels (i.e., Levels 1, 2, 3, or 4) within each grade-band and content area assessment. The correspondence between reported (TA) performance levels and VR (or SR) performance levels was assessed according to the kappa and weighted kappa coefficients. In particular, consistency was assessed through the *weighted kappa statistic* (Agresti, 1990; Spitzer, Cohen, Fleiss, & Endicott, 1967), which is appropriate for ordered categories:

$$\kappa_w = \frac{\sum \sum w_{ij} \pi_{ij} - \sum \sum w_{ij} \pi_{i+} \pi_{+j}}{1 - \sum \sum w_{ij} \pi_{i+} \pi_{+j}},$$

where i is the category assigned by the TA, j is the category assigned by the VR, $w_{ij} = 1 - (i - j)^2 / (I - 1)^2$ are the weights, π_{ij} is the probability of being classified as ij , and “+” indicates agreement between categories. Kappa equals 0 when the agreement is that expected by chance; and kappa equals 1 when there is perfect agreement among raters.

Under the current condition, it must be noted that not all cases included in this analysis contained complete data. The “N” rows of Exhibit C-4 indicate the *effective sample size* (where “n” is the count of valid TA administrations with complete VR or SR items scores). For these realized samples, there is a high level of agreement: The weighted kappa coefficients range between 0.91 and 0.99, and their 95% confidence intervals fall within ± 0.061 of the point estimates.

Exhibit C-4: Agreement Statistics by Method, Subject, and Grade-Band

Subject	Statistic	Second Rater	Video Rater		
		Elementary	Elementary	Middle	High
ELA	N	48	70	61	44
	<i>kw</i>	0.957	0.905	0.934	0.988
	95% CI	0.916–0.998	0.851–0.959	0.889–0.979	0.964–1
Biology	N				41
	<i>kw</i>				0.927
	95% CI				0.866–0.987

Summary

TA and VR assignments of students to performance levels typically show high levels of agreement, as weighted kappa typically ranges from 0.91 to 1.00. Further, the 95% confidence intervals show that, while sample sizes for the current calculations may be small, the agreement indices are significantly greater than chance agreement and often approach 1.00. Based on the current evidence, we can conclude that the SC-Alt was accurately scored and that the second rater and video study approaches yield similar results.

Appendix D: Descriptions of Achievement Levels (DALs)

Exhibit D-1: English Language Arts Descriptions of Achievement Levels

Performance Level	ELA Achievement Level Definitions	Grade-Band 3–5	Grade-Band 6–8	Grade 10
1	Students performing at Level 1 demonstrate emerging academic skills and competencies in reading, writing, and communication.	<p>Students performing at Level 1 should be able to</p> <ul style="list-style-type: none"> listen (as evidenced by facial expressions, gestures, or sounds) to a variety of text read aloud; point or eye gaze to objects, pictures, or letters to complete a writing activity; engage (using facial expressions, gestures, or sounds) in conversations focused on objects in the immediate surroundings; listen (as evidenced by facial expressions, gestures, or sounds) to a speaker. 	<p>Students performing at Level 1 should be able to</p> <ul style="list-style-type: none"> point or eye gaze to objects or pictures related to a variety of grade-appropriate or adapted text focused on concrete concepts, read aloud; point or eye gaze to objects, pictures, or letters to create a simple composition; engage in conversations focused on events in the immediate surroundings as evidenced by facial expressions, gestures, or sounds; listen to a speaker as evidenced by facial expressions or gestures without interrupting. 	<p>Students performing at Level 1 should be able to</p> <ul style="list-style-type: none"> listen to a variety of grade-appropriate/adapted texts read aloud as evidenced by facial expressions, gestures, or sounds; point or eye gaze to objects, pictures, or letters to complete more complex written products; engage in conversations focused on objects or events outside the immediate surroundings as evidenced by facial expressions, gestures, or sounds; listen and respond to a speaker.
2	Students performing at Level 2 demonstrate foundational academic skills and competencies in reading, writing, and communication.	<p>Students performing at Level 2 should be able to</p> <ul style="list-style-type: none"> tell or show what a grade-appropriate or adapted text, which contains high-frequency words, is about; identify individual words/picture symbols; identify story elements (e.g., main idea, events, 	<p>Students performing at Level 2 should be able to</p> <ul style="list-style-type: none"> tell or show what a text that requires only literal interpretation is about (using objects, pictures, or words); read a variety of grade-appropriate/adapted texts (e.g., recipes or advertisements); identify story elements (e.g., 	<p>Students performing at Level 2 should be able to</p> <ul style="list-style-type: none"> tell or show what a grade-appropriate or adapted text that requires simple inferences is about; read a variety of texts (e.g., recipes, advertisements, schedules, and newspapers); identify story elements (e.g., main idea, events, setting, characters,

Performance Level	ELA Achievement Level Definitions	Grade-Band 3–5	Grade-Band 6–8	Grade 10
		setting, and characters); <ul style="list-style-type: none"> • use oral and written language to describe; • select from a list of topics to generate ideas for written communication; • listen to a speaker without interrupting; • respond appropriately in conversations. 	main idea, events, setting, characters, and conflict); <ul style="list-style-type: none"> • make connections within and between texts; • use oral and written language to explain; • select from a list of topics to generate multiple ideas for written communication; • focus attention on a speaker and listen without interrupting; • engage in conversations by answering direct questions about familiar situations; • follow oral and/or written one-step directions. 	conflict, and plot); <ul style="list-style-type: none"> • gather meaning from graphic representations; • use oral and written language to explain, inform, and describe; • generate ideas for written communication; • edit own writing; • engage in conversations by answering direct questions about the immediate environment or other familiar surroundings.
3	Students performing at Level 3 demonstrate increasing academic skills and competencies in reading, writing, and communication.	Students performing at Level 3 should be able to <ul style="list-style-type: none"> • identify story elements in text (e.g., characters, settings, events, cause and effect, and problem and solution); • read words and simple sentences; • generate an idea and use words, pictures, or oral language to write; • follow one-step oral or signed directions; • communicate agreement or disagreement appropriately. 	Students performing at Level 3 should be able to <ul style="list-style-type: none"> • identify and recall details in text including main idea and characters; • draw conclusions and make simple predictions and inferences about the text; • determine meaning of unfamiliar words; • generate multiple ideas by selecting from a list and use words, pictures, or oral language to write; • initiate conversation about immediate surroundings. 	Students performing at Level 3 should be able to <ul style="list-style-type: none"> • make connections with text (plot, characters, setting); • make inferences about events in text; • understand multiple meanings of words; • compare and contrast story elements from different stories; • discriminate fact from fiction; • generate an idea and use words, pictures, or oral language to write; • follow directions to complete a task; • initiate conversations about immediate surroundings or other familiar topics.

Performance Level	ELA Achievement Level Definitions	Grade-Band 3–5	Grade-Band 6–8	Grade 10
4	Students performing at Level 4 demonstrate and apply academic skills and competencies in reading, writing, and communication.	<p>Students performing at Level 4 should be able to</p> <ul style="list-style-type: none"> • identify story elements such as the main idea and cause and effect; • draw conclusions and make predictions about text; • read and understand the main idea of a simple paragraph; • create and edit personal written products; • follow two-step oral or signed directions; • take turns appropriately during conversation or discussion. 	<p>Students performing at Level 4 should be able to</p> <ul style="list-style-type: none"> • recognize and recall details in text, including the main idea, plot, characters, and setting; • draw conclusions and make predictions and inferences about the text; • read and understand the main idea of a simple paragraph; • explain word meanings; • create and edit personal written products; • follow oral/signed or written directions; • initiate and retell conversations. 	<p>Students performing at Level 4 should be able to</p> <ul style="list-style-type: none"> • recognize and recall details in text, including the main idea, plot, characters, and setting; • draw conclusions and make predictions and inferences about the text; • read and understand the main idea of a short story; • use context clues to understand the meaning of unknown words; • make connections within and between texts and to prior knowledge, other texts, and the world; • create and edit personal written products; • use graphic representations as sources of information.

Exhibit D-2: Mathematics Descriptions of Achievement Levels

Performance Level	Mathematics Achievement Level Definitions	Grade-Band 3–5	Grade-Band 6–8	Grade 10
1	Students performing at Level 1 demonstrate emerging academic skills and competencies in mathematics.	<p>Students performing at Level 1 should be able to</p> <ul style="list-style-type: none"> manipulate one concrete object; observe that two geometric figures have the same attributes; recognize attributes of objects, such as length and weight. 	<p>Students performing at Level 1 should be able to</p> <ul style="list-style-type: none"> recognize the concept of one in counting objects; recognize that two geometric figures have the same attributes; recognize attributes of objects, such as length, weight, and size/volume. 	<p>Students performing at Level 1 should be able to</p> <ul style="list-style-type: none"> recognize the concept of one more in counting objects; match geometric figures that have the same attributes; respond to positional concepts such as on top of/under, on/off, above/below; match objects by one attribute such as length, weight, and size/volume.
2	Students performing at Level 2 demonstrate foundational academic skills and competencies in mathematics.	<p>Students performing at Level 2 should be able to</p> <ul style="list-style-type: none"> count objects in a set; identify objects by one attribute (color, size, shape); classify two - and three-dimensional concrete objects according to one attribute; recognize positional concepts (on/off); identify measurement tools, including graphs. 	<p>Students performing at Level 2 should be able to</p> <ul style="list-style-type: none"> add and subtract using concrete objects; sort objects by one attribute (color, size, shape); recognize and demonstrate understanding of positional concepts (on/off, below/above); use nonstandard units to measure; match the correct tool to a specific task (e.g., measure length, weight, time); identify parts of a chart, graph, or table. 	<p>Students performing at Level 2 should be able to</p> <ul style="list-style-type: none"> solve addition and subtraction problems; Identify operations (+ or -); tell which has more in a set; identify a repeating relationship (pattern); sort and classify objects by one attribute (length, height, weight, volume); use a graph or chart to gain information.
3	Students performing at	Students performing at Level 3	Students performing at Level 3	Students performing at Level 3

Performance Level	Mathematics Achievement Level Definitions	Grade-Band 3–5	Grade-Band 6–8	Grade 10
	Level 3 demonstrate increasing academic skills and competencies in mathematics.	<p>should be able to</p> <ul style="list-style-type: none"> • demonstrate addition and subtraction concretely or symbolically; • count and compare objects in a set; • sort and classify objects by attribute (shape, size); • identify three-dimensional shapes (cube, sphere, cylinder); • use nonstandard units to measure; • find answers to questions in a graph. 	<p>should be able to</p> <ul style="list-style-type: none"> • identify the answer to one-digit addition and subtraction problems; • identify a set as having more, fewer, or the same number as another set; • extend a repeating pattern; • compare objects by attribute; • interpret information displayed in a table. 	<p>should be able to</p> <ul style="list-style-type: none"> • identify the process for solving an addition or a subtraction problem; • identify and use operational symbols correctly; • estimate the number of objects in a set; • add to find value of a set of coins; • describe, create, and complete a repeating pattern; • use and organize data to create charts, graphs, and tables.
4	Students performing at Level 4 demonstrate and apply academic skills and competencies in mathematics.	<p>Students performing at Level 4 should be able to</p> <ul style="list-style-type: none"> • demonstrate understanding of addition and subtraction; • generate a pattern using three-dimensional shapes (cube, sphere, cylinder); • compare objects by attribute (length, size); • interpret information displayed in a graph. 	<p>Students performing at Level 4 should be able to</p> <ul style="list-style-type: none"> • solve addition and subtraction facts without regrouping; • describe and extend a repeating pattern; • interpret information displayed in a graph; • use data to create tables. 	<p>Students performing at Level 4 should be able to</p> <ul style="list-style-type: none"> • identify, compare, and construct numbers; • use operation symbols (more than, less than, and equal to) to solve problems; • add to find the value of a set of two or more coins; • identify, describe, create, extend, and complete a repeating pattern; • describe events as more likely or less likely to occur; • use and organize data to create and interpret graphs.

Exhibit D-3: Science Descriptions of Achievement Levels

Performance Level	Science Achievement Level Definitions	Grade-Band 3–5	Grade-Band 6–8	Grade 10
1	Students performing at Level 1 demonstrate emerging academic skills and competencies in science.	<p>Students performing at Level 1 should be able to use their senses to</p> <ul style="list-style-type: none"> observe the outcome of a simple science investigation; sequence growth patterns; observe and record daily weather conditions; recognize the sun and moon and relate them to day and night; recognize that objects move when force is applied. 	<p>Students performing at Level 1 should be able to use their senses to</p> <ul style="list-style-type: none"> choose a question (how) (what if) to conduct a scientific investigation; identify major body parts of animals; identify the sun and moon; recognize that objects move when force is applied and recognize speed (fast and slow); sort by one attribute. 	<p>Students performing at Level 1 should be able to use their senses to</p> <ul style="list-style-type: none"> choose questions to conduct a simple scientific investigation; recognize that objects move when force is applied; recognize that an object at rest does not move; identify physical properties of matter (e.g., freezing/melting)
2	Students performing at Level 2 demonstrate foundational academic skills and competencies in science.	<p>Students performing at Level 2 should be able to</p> <ul style="list-style-type: none"> generate a question to conduct a simple scientific investigation; sort organisms by physical characteristics; identify daily weather conditions; recognize the pattern of day and night; identify the position of objects such as above/below, inside, or on top; sort materials by observable properties. 	<p>Students performing at Level 2 should be able to</p> <ul style="list-style-type: none"> carry out a simple scientific investigation to answer a question; sort and describe materials by observable properties; sort and identify organisms by physical characteristics; identify patterns of day and night; recognize that an object at rest moves when force is applied. 	<p>Students performing at Level 2 should be able to</p> <ul style="list-style-type: none"> carry out a simple scientific investigation related to electricity or force and motion to answer a question; compare magnetic and non-magnetic objects; identify the force that makes an object move; recognize physical changes in matter; recognize physical properties of matter.
3	Students performing at	Students performing at Level 3 should be able to	Students performing at Level 3 should be able to	Students performing at Level 3 should be able to

Performance Level	Science Achievement Level Definitions	Grade-Band 3–5	Grade-Band 6–8	Grade 10
	Level 3 demonstrate increasing academic skills and competencies in science.	<ul style="list-style-type: none"> select appropriate tool for gathering data; carry out a simple scientific investigation; classify events in sequential order; distinguish between living and nonliving things; identify major organs of animals; use a graph to compare daily changes in weather conditions. 	<ul style="list-style-type: none"> conduct and analyze the results of a simple scientific investigation; use graphs, tables, and charts to record data and report on the results of an investigation; compare the characteristics of living and nonliving things; identify what plants need to grow; use a graph or chart to compare weather conditions each season; classify organism into major groups. 	<ul style="list-style-type: none"> predict the outcome of a simple investigation and compare the results with the prediction; compare factors that affect an electromagnet; identify electricity as a source of energy; relate the change in force to the change in speed; recognize the physical properties of two or more objects.
4	Students performing at Level 4 demonstrate and apply academic skills and competencies in science.	<p>Students performing at Level 4 should</p> <ul style="list-style-type: none"> plan and conduct a simple scientific investigation; identify major organs of animals and their functions; identify living and nonliving things in terms of a food web; identify natural resources as renewable or nonrenewable; compare heat and light changes from season to season using a graph; draw simple conclusions from tables, graphs, and charts 	<p>Students performing at Level 4 should be able to</p> <ul style="list-style-type: none"> plan, conduct, and carry out a simple scientific investigation; communicate simple conclusions using tables and graphs; identify simple machines (inclined plane, lever, pulley); compare data on temperature changes over time using a graph; use a graph to show how heat and light change from season to season; identify sources of light. 	<p>Students performing at Level 4 should be able to</p> <ul style="list-style-type: none"> plan, conduct, and analyze the results of a scientific investigation; draw simple conclusions from distance/time graphs or tables; demonstrate how simple machines are used to help people (inclined plane, lever, pulley, etc.); predict the effect of the change in force on an object; identify water as solid, steam, or liquid.

Exhibit D-4: Social Studies Descriptions of Achievement Levels

Performance Level	Social Studies Achievement Level Definitions	Grade-Band 3–5	Grade-Band 6–8
1	Students performing at Level 1 demonstrate emerging academic skills and competencies in social studies.	<p>Students performing at Level 1 should be able to</p> <ul style="list-style-type: none"> • identify self from others; • respond to a person in authority in the home or school; • follow class rules; • engage in turn-taking; • listen to information about South Carolina history. 	<p>Students performing at Level 1 should be able to</p> <ul style="list-style-type: none"> • identify self from others; • respond to familiar authority figures; • follow class rules; • engage in turn-taking and sharing; • listen to information presented about significant and historical events in South Carolina.
2	Students performing at Level 2 demonstrate foundational skills and competencies in social studies.	<p>Students performing at Level 2 should be able to</p> <ul style="list-style-type: none"> • identify characteristics such as gender that help identify self in relation to others; • match workers to different jobs in the community; • recognize people in authority and follow class rules; • match the people we honor on some national holidays (e.g., George Washington, Martin Luther King, Jr.) with the holidays; • distinguish between past and present (match jobs of the past with jobs of the present); • match significant historical figures such as Thomas Edison to their accomplishments. 	<p>Students performing at Level 2 should be able to</p> <ul style="list-style-type: none"> • identify surroundings (e.g., classroom, school); • match different people to their jobs in the community; • identify people in authority and follow class rules; • demonstrate understanding of rules; • identify the people we honor on some national holidays (e.g., George Washington, Martin Luther King, Jr.); • identify the purpose of money; • match changes over time to the past and present such as communication.
3	Students performing at Level 3 demonstrate increasing skills and competencies in social studies.	<p>Students performing at Level 3 should be able to</p> <ul style="list-style-type: none"> • understand the concept of past and present; • demonstrate respect for people in authority; • identify major symbols of the United States; • identify why we celebrate the national holidays; • recognize that when we work we earn money 	<p>Students performing at Level 3 should be able to</p> <ul style="list-style-type: none"> • identify members of the larger community (e.g., police officers, firefighters, doctors); • demonstrate understanding of the consequences of not following the rules; • identify examples of good citizenship such as honesty, courage, etc.; • identify symbols of the United States (e.g., the flag, bald

Performance Level	Social Studies Achievement Level Definitions	Grade-Band 3–5	Grade-Band 6–8
		<p>to buy things;</p> <ul style="list-style-type: none"> identify features on a map of South Carolina (river, mountain, ocean); answer questions about significant events related to the Civil War; identify historical figures such as Thomas Edison, Alexander Graham Bell, etc., to their accomplishments. 	<p>eagle);</p> <ul style="list-style-type: none"> demonstrate an understanding that we work to earn money and use money to buy things; identify changes over time such as in travel, farming, etc.; gain information from maps, charts, and graphs; answer questions about key historical figures and significant historical events including the civil rights movement.
4	Students performing at Level 4 demonstrate and apply academic skills and competencies in social studies.	<p>Students performing at Level 4 should be able to</p> <ul style="list-style-type: none"> place personal history on a time line; identify the roles of leaders and officials in local government (e.g., principal, mayor, governor); identify individuals who embody qualities of good citizenship; identify examples of respect and fair treatment; recognize that we exchange money for goods and services; use a key to locate geographic features on a map of South Carolina; answer questions about key concepts related to the Civil War; answer questions about the accomplishments of key historical figures such as Thomas Edison, Alexander Graham Bell, etc. 	<p>Students performing at Level 4 should be able to</p> <ul style="list-style-type: none"> place personal and family history on a time line; identify roles of leaders and officials in local government (e.g., principal, mayor, governor); identify examples of the qualities of courage and patriotism; identify examples of respect and fair treatment and their opposites; recognize how the amount of money available determines what we can buy; gain information from maps and charts; identify the accomplishments of Civil Rights leaders including Rosa Parks.

Exhibit D-5: High School Biology Descriptions of Achievement Levels

Performance Level	Biology Achievement Level Definitions	Grade 10
1	Students performing at Level 1 demonstrate some emerging academic skills and competencies in biology.	<p>Students performing at Level 1 should be able to</p> <ul style="list-style-type: none"> • Identify a possible outcome of a simple scientific investigation • Recognize tools that could be used in a simple scientific investigation • Identify a result of a simple investigation based on observations • Identify appropriate safety instruments when conducting scientific investigations • Identify things as cellular (living) • Identify food as a source of protein, carbohydrates, or fat • Identify the source of energy in a food chain • Identify the offspring of parents • Identify adaptations that allow animals to survive in their habitat • Identify living and nonliving resources in an ecosystem • Identify natural things in the environment and things made by humans
2	Students performing at Level 2 demonstrate foundational academic skills in biology.	<p>Students performing at Level 2 should be able to</p> <ul style="list-style-type: none"> • Identify a prediction • Identify the outcome of a simple controlled scientific investigation • Identify scientific instruments used to make observations • Interpret simple scientific data • Identify parts of a graph • Identify appropriate safety procedures when conducting scientific investigations • Recognize cellular vs. non cellular (living or nonliving) things • Recognize food as protein, carbohydrate, or fat • Identify the flow of energy in a simple food web • Identify parents as a source of physical traits • Identify favorable and unfavorable traits that determine species survival • Identify a phylogenetic tree as a diagram that shows ancestry of organisms • Recognize the relationships among organisms • Identify environmental changes that can effect a population • Identify human activities that affect Earth

Performance Level	Biology Achievement Level Definitions	Grade 10
3	Students performing at Level 3 demonstrate increasing academic skills and competencies in biology.	<p>Students performing at Level 3 should be able to</p> <ul style="list-style-type: none"> • Identify the hypothesis of a simple investigation • Recognize which scientific instruments are used to collect and/or record data • Organize data in a given graph/table/model • Interpret the results of scientific data that are displayed in a graph • Identify the outcome of a simple investigation as the same/different from the original hypothesis • Identify appropriate safety procedures required when conducting a specific scientific investigation • Recall that cells are the basic unit of life • Classify things as cellular or non cellular • Illustrate that all living things are composed of cells • Identify different types of cells, tissues, and organs • Illustrate the end product of cell division • Identify what plants need for survival • Classify different foods as protein, fat, or carbohydrate • Summarize the role of protein, carbohydrates, or fat on the body • Illustrate the flow of energy in a simple food web • Identify that chromosomes contain DNA • Identify types of traits passed on from parent to offspring • Identify offspring based on dominant parent traits • Identify the structure of DNA • Identify an organism that is better adapted to a changing habitat • Identify which organisms are most closely related by using a phylogenetic tree • Identify predator/prey relationships • Explain how environmental changes can affect a population • Identify the sequence of ecological succession • Classify human activities based on their effect on Earth (beneficial or harmful)

Performance Level	Biology Achievement Level Definitions	Grade 10
4	Students performing at Level 4 demonstrate and apply academic skills in biology.	<p>Students performing at Level 4 should be able to</p> <ul style="list-style-type: none"> Analyze the outcome of a simple investigation and compare it to the hypothesis Select the appropriate graph for displaying simple scientific data Use laboratory instruments and procedures in a safe manner Recall that all cells come from other cells Identify a nucleus, cell membrane/wall, vacuole, and chloroplast Recall different types of cells Illustrate that plants and animals have different cell structures Identify different types of cells, tissues, organs, and organ systems Classify protein, carbohydrate, or fats based on function or description of structure Create a food web showing the flow of energy Summarize that plants use photosynthesis to make their own food Identify that DNA and genes pass on specific traits to offspring Predict physical traits of offspring based on dominant or recessive physical traits of parents Identify a dominant trait of a given species Identify the principle of natural selection Explain the effect of a changing habitat on a population Explain the relationship of two organisms based on a phylogenetic tree Identify living counterparts of extinct organisms Classify interrelationships among organisms within ecosystems Predict the effect of environmental changes on a population Illustrate the changes that occur during succession Illustrate how human activities affect the naturally occurring processes on Earth

Appendix E: Summary of Linking Design

Exhibit E-1: Summary of Linking Design across Subjects and Grade-Bands, for ELA, Math, Science, and Social Studies

Subject	Grade-Band	Number of Items	Number of Tasks	Starting Positions									
				Starting Task 1		Within Grade-Band Linking		Starting Task 3		Within Grade-Band Linking		Starting Task 7	
				Items	Tasks	Items	Tasks	Items	Tasks	Items	Tasks	Items	Tasks
ELA	3–5	83	15	39	7	29	5	50	9	26	5	49	9
	(Linking)	16	3										
	6–8	79	15	38	7	26	5	48	9	26	5	45	9
	(Linking)	23	5										
	10	80	15	36	7	24	5	45	9	26	5	49	9
Math	3–5	76	15	38	7	28	5	46	9	25	5	45	9
	(Linking)	33	7										
	6–8	75	15	33	7	23	5	45	9	26	5	46	9
	(Linking)	20	4										
	10	81	15	34	7	25	5	50	9	30	5	52	9
Science	3–5	77	15	35	7	25	5	48	9	27	5	46	9
	(Linking)	-	-										
	6–8	79	15	40	7	29	5	49	9	25	5	44	9
Social Studies	3–5	80	15	32	7	24	5	47	9	27	5	52	9
	(Linking)	16	4										
	6–8	71	15	29	7	21	5	41	9	24	5	46	9

Exhibit E-2: Summary of Linking Design, High School Biology

Subject	Grade	Number of Items	Number of Tasks	Starting Positions									
				Starting Task 1		Within Grade-Band Linking		Starting Task 3		Within Grade-Band Linking		Starting Task 6	
				Items	Tasks	Items	Tasks	Items	Tasks	Items	Tasks	Items	Tasks
Biology	10	65	12	33	6	22	4	40	7	24	4	38	7

Appendix F: Statistics Summaries for the 2011 Spring Field-Test Items

Exhibit F-1: Grade-Band 3-5 ELA Field-Test Classical Item Statistics

ITS Item ID	Grade	Item Position	Adjusted Biserial/ Polyserial	Average Score	Access Limitation	Omit	DIF	
							Female vs. Male	Black vs. White
1770	3-5	23	0.36	0.4555385	0.00	0.03	-A	-A
1771	3-5	24	0.45	0.648119	0.00	0.02	+A	+A
1767	3-5	25	0.32	0.398744	0.00	0.02	-B	-A
1772	3-5	26	0.31	0.3492935	0.00	0.03	-C	+A
1768	3-5	27	0.44	0.351525	0.01	0.03	-A	-A
1769	3-5	28	0.37	0.4843015	0.00	0.03	+A	-A
1778	3-5	40	0.48	0.6640565	0.00	0.01	-A	+A
1774	3-5	41	0.22	0.540329	0.00	0.01	+A	-A
1775	3-5	42	0.36	0.678152	0.00	0.01	-A	+A
1776	3-5	43	0.31	0.5101885	0.00	0.01	-A	-A
1777	3-5	44	0.44	0.769835	0.00	0.01	+A	+A
1773	3-5	45	0.30	0.443396	0.00	0.02	-A	-A
1706	3-5	61	0.58	0.605566	0.00	0.00	+A	+A
1707	3-5	62	0.55	0.732246	0.00	0.00	+A	+A
1708	3-5	63	0.24	0.6517775	0.00	0.00	-A	-B
1709	3-5	64	0.40	0.509126	0.00	0.00	-A	-A
1710	3-5	65	0.52	0.5846155	0.00	0.00	+A	-A
1711	3-5	66	0.36	0.7949955	0.00	0.01	+A	+A

Exhibit F-2: Grade-Band 6-8 ELA Field-Test Classical Item Statistics

ITS Item ID	Grade	Item Position	Adjusted Biserial/ Polyserial	Average Score	Access Limitation	Omit	DIF	
							Female vs. Male	Black vs. White
1784	6-8	24	0.44	0.638554	0.00	0.02	-A	+A
1779	6-8	25	0.39	0.408081	0.00	0.03	-A	-A
1780	6-8	26	0.43	0.468687	0.00	0.02	-A	+A
1782	6-8	27	0.27	0.3838385	0.00	0.03	-A	+A
1783	6-8	28	0.43	0.6123485	0.00	0.03	+A	-A
1712	6-8	39	0.57	0.546902	0.00	0.00	+A	+A
1713	6-8	40	0.54	0.7056035	0.00	0.01	+A	+A
1714	6-8	41	0.63	0.512931	0.00	0.01	+A	+A
1715	6-8	42	0.63	0.75475	0.00	0.00	-A	-A
1717	6-8	43	0.45	0.704408	0.00	0.01	+A	-A
1716	6-8	44	0.12	0.4956745	0.00	0.01	-A	+A
1830	6-8	61	0.29	0.5863205	0.00	0.00	+A	+A
1831	6-8	62	0.62	0.5863205	0.00	0.00	+A	+A
1832	6-8	63	0.65	0.534434	0.00	0.00	-A	-A
1833	6-8	64	0.25	0.4191875	0.00	0.01	+A	+A
1834	6-8	65	0.32	0.632798	0.00	0.01	+A	+A
1835	6-8	66	0.41	0.6471145	0.00	0.01	+A	+A

Exhibit F-3: Grade 10 ELA Field-Test Classical Item Statistics

ITS Item ID	Grade	Item Position	Adjusted Biserial/ Polyserial	Average Score	Access Limitation	Omit	DIF	
							Female vs. Male	Black vs. White
1804	10	22	0.40	0.548872	0.00	0.01	+A	+A
1806	10	23	0.45	0.473684	0.00	0.01	-A	+A
1813	10	24	0.49	0.676692	0.00	0.02	-A	+A
1808	10	25	0.62	0.5112785	0.00	0.01	+A	+A
1810	10	26	0.54	0.5338345	0.00	0.02	+A	+A
1766	10	37	0.72	0.823052	0.00	0.01	+A	-A
1823	10	38	0.71	0.8295455	0.00	0.01	-A	-A
1825	10	39	0.64	0.5811685	0.00	0.01	-A	+A
1826	10	40	0.55	0.75	0.00	0.00	-A	+B
1827	10	41	0.46	0.7483765	0.00	0.01	-A	+A
1803	10	58	0.56	0.739927	0.00	0.00	+A	-A
1800	10	59	0.49	0.64652	0.00	0.00	-A	-A
1801	10	60	0.32	0.3882785	0.00	0.00	+A	+A
1802	10	61	0.33	0.6428575	0.00	0.00	+A	-A
1799	10	62	0.59	0.610701	0.00	0.01	+A	-A
1798	10	63	0.28	0.5055355	0.00	0.01	-A	+A

Exhibit F-4: Grade-Band 3-5 Math Field-Test Classical Item Statistics

ITS Item ID	Grade	Item Position	Adjusted Biserial/ Polyserial	Average Score	Access Limitation	Omit	DIF	
							Female vs. Male	Black vs. White
1634	3-5	22	0.74	0.7007573	0.00	0.01	+A	-A
1635	3-5	23	0.47	0.516595	0.00	0.04	-A	+A
1636	3-5	24	0.36	0.645441	0.00	0.04	+A	-A
1637	3-5	25	0.41	0.398256	0.00	0.04	-A	+A
1638	3-5	26	0.51	0.533479	0.00	0.05	-A	-A
1639	3-5	27	0.44	0.517493	0.00	0.05	-A	-A
1646	3-5	39	0.43	0.5614245	0.00	0.02	+A	-A
1647	3-5	40	0.46	0.627934	0.00	0.01	+A	-A
1648	3-5	41	0.47	0.567398	0.00	0.02	-A	-A
1649	3-5	42	0.52	0.522763	0.00	0.01	-A	+A
1650	3-5	43	0.52	0.6415615	0.00	0.02	+A	+A
1651	3-5	44	0.48	0.602534	0.00	0.02	-A	+A
1640	3-5	57	0.08	0.335255	0.00	0.00	+A	-A
1641	3-5	58	0.36	0.716763	0.00	0.01	-A	-A
1642	3-5	59	0.50	0.756027	0.00	0.01	+A	+A
1643	3-5	60	0.39	0.5737705	0.00	0.01	-A	-A
1644	3-5	61	0.29	0.650916	0.00	0.01	+A	-A
1645	3-5	62	0.58	0.7193235	0.00	0.01	+A	+A

Exhibit F-5: Grade-Band 6-8 Math Field-Test Classical Item Statistics

ITS Item ID	Grade	Item Position	Adjusted Biserial/ Polyserial	Average Score	Access Limitation	Omit	DIF	
							Female vs. Male	Black vs. White
1730	6-8	20	0.39	0.4362345	0.00	0.02	-A	+B
1731	6-8	21	0.46	0.612576	0.00	0.03	-A	+A
1732	6-8	22	0.34	0.38641	0.00	0.03	-A	+A
1733	6-8	23	0.47	0.520325	0.00	0.03	+A	-A
1734	6-8	24	0.43	0.320122	0.00	0.04	+A	-A
1735	6-8	25	0.37	0.422764	0.00	0.03	+A	+A
1718	6-8	34	0.36	0.475171	0.00	0.01	-A	+A
1719	6-8	35	0.41	0.519726	0.00	0.01	-A	-A
1722	6-8	36	0.25	0.49742	0.00	0.01	-A	+A
1723	6-8	37	0.43	0.4952585	0.00	0.01	+A	+A
1724	6-8	56	0.22	0.6171345	0.00	0.00	+A	-A
1726	6-8	57	0.26	0.586641	0.00	0.01	-A	-A
1725	6-8	58	0.40	0.5891475	0.00	0.01	-A	+A
1727	6-8	59	0.48	0.483981	0.00	0.01	-A	+A
1728	6-8	60	0.33	0.5967895	0.00	0.01	+A	-A
1729	6-8	61	0.26	0.531615	0.00	0.01	-A	+A

Exhibit F-6: Grade 10 Math Field-Test Classical Item Statistics

ITS Item ID	Grade	Item Position	Adjusted Biserial/ Polyserial	Average Score	Access Limitation	Omit	DIF	
							Female vs. Male	Black vs. White
1759	10	20	0.79	0.6833333	0.00	0.03	+A	-A
1761	10	21	0.42	0.448148	0.00	0.04	-A	-A
1762	10	22	0.48	0.4552235	0.00	0.04	-A	-A
1763	10	23	0.34	0.4291045	0.00	0.04	+A	-A
1764	10	24	0.47	0.593284	0.00	0.04	+A	+A
1752	10	35	0.31	0.528053	0.00	0.01	+A	-A
1753	10	36	0.24	0.5346535	0.00	0.01	-A	-A
1754	10	37	0.44	0.70297	0.00	0.00	-B	+A
1755	10	38	0.45	0.5627065	0.00	0.00	+A	+A
1756	10	39	0.41	0.5676565	0.00	0.01	+A	-A
1757	10	40	0.21	0.537954	0.00	0.01	+A	-A
1758	10	41	0.19	0.4554455	0.00	0.01	-A	+A
1742	10	60	0.43	0.612782	0.00	0.00	+A	-A
1743	10	61	0.32	0.6954885	0.00	0.00	-A	+A
1744	10	62	0.40	0.4454885	0.00	0.01	+A	+A
1745	10	63	0.44	0.733083	0.00	0.01	+A	-A
1746	10	64	0.48	0.8421055	0.00	0.00	-A	+C

Exhibit F-7: Grade-Band 3-5 Science Field-Test Classical Item Statistics

ITS Item ID	Grade	Item Position	Adjusted Biserial/ Polyserial	Average Score	Access Limitation	Omit	DIF	
							Female vs. Male	Black vs. White
894	3-5	11	0.77	0.7871797	0.00	0.00	-A	+A
1023	3-5	12	0.64	0.6964980	0.00	0.03	-A	-A
895	3-5	13	0.58	0.6530210	0.00	0.04	-A	+A
896	3-5	14	0.49	0.5605470	0.00	0.04	-A	-A
898	3-5	15	0.49	0.5166670	0.00	0.05	-A	+A
965	3-5	16	0.52	0.6907630	0.00	0.02	+A	-A
966	3-5	17	0.61	0.6726905	0.00	0.01	-A	+A
967	3-5	18	0.42	0.6365460	0.00	0.02	+A	-A
968	3-5	19	0.50	0.5171030	0.00	0.02	-A	-A
969	3-5	20	0.47	0.5947580	0.00	0.03	+A	-A
1663	3-5	21	0.57	0.6566265	0.00	0.02	-A	+A
1658	3-5	22	0.50	0.5823295	0.00	0.02	+A	-A
1659	3-5	23	0.53	0.5823290	0.00	0.02	-A	-A
1660	3-5	24	0.58	0.6428570	0.00	0.02	-A	-A
1661	3-5	25	0.52	0.5010060	0.00	0.02	-A	-A
1662	3-5	26	0.45	0.5443545	0.00	0.03	+A	-A
1669	3-5	36	0.39	0.7100460	0.00	0.00	+A	+A
1665	3-5	37	0.65	0.8167805	0.00	0.00	+A	+A
1828	3-5	38	0.67	0.6877850	0.00	0.00	-A	+A
1666	3-5	39	0.66	0.7899545	0.00	0.00	+A	+A
1829	3-5	40	0.55	0.4902860	0.00	0.00	-A	+A
1668	3-5	41	0.37	0.6571100	0.00	0.01	+A	-A
1670	3-5	59	0.48	0.7503825	0.00	0.00	+A	+A
1673	3-5	60	0.58	0.7346625	0.00	0.00	+A	+A
1675	3-5	61	0.31	0.7811060	0.00	0.00	+A	-A
1672	3-5	62	0.47	0.5546995	0.00	0.01	-A	+A
1671	3-5	63	0.19	0.6833590	0.00	0.00	+A	-A
1674	3-5	64	0.12	0.4476120	0.00	0.00	-A	-A

Exhibit F-8: Grade-Band 6-8 Science Field-Test Classical Item Statistics

ITS Item ID	Grade	Item Position	Adjusted Biserial/ Polyserial	Average Score	Access Limitation	Omit	DIF	
							Female vs. Male	Black vs. White
1676	6-8	24	0.55	0.6309225	0.00	0.03	-A	+A
1677	6-8	25	0.41	0.6084790	0.00	0.02	-A	+A
1678	6-8	26	0.53	0.4551125	0.00	0.03	-A	-A
1679	6-8	27	0.36	0.4762500	0.00	0.02	-A	-A
1680	6-8	28	0.29	0.4210525	0.00	0.03	-A	+A
1681	6-8	29	0.41	0.3634085	0.00	0.04	-A	+A
957	6-8	36	0.52	0.8281250	0.00	0.01	-A	+A
959	6-8	37	0.75	0.6640535	0.00	0.03	-A	+A
958	6-8	38	0.68	0.6813060	0.01	0.02	-A	+A
960	6-8	39	0.65	0.7448865	0.01	0.03	-A	-A
963	6-8	40	0.65	0.7443310	0.00	0.02	-A	-A
1682	6-8	41	0.63	0.7349625	0.00	0.00	-A	-A
1683	6-8	42	0.63	0.7233375	0.00	0.00	-A	-A
1685	6-8	43	0.28	0.4911950	0.00	0.01	+A	-A
1686	6-8	44	0.59	0.5018870	0.00	0.01	+A	-A
1696	6-8	61	0.73	0.7481965	0.00	0.00	+A	-A
1694	6-8	62	0.41	0.6969695	0.00	0.00	+A	-A
1695	6-8	63	0.54	0.7582975	0.00	0.00	+A	-A
1697	6-8	64	0.59	0.6486295	0.00	0.01	+A	+A
1698	6-8	65	0.44	0.5743145	0.00	0.00	-A	+A
1699	6-8	66	0.50	0.6589595	0.00	0.01	-A	+A
1003	6-8	76	0.61	0.7716050	0.00	0.00	+A	+A
1004	6-8	77	0.04	0.3769350	0.00	0.00	-A	-A
1005	6-8	78	0.52	0.5046300	0.00	0.00	+A	+A
1006	6-8	79	0.47	0.5409580	0.00	0.00	+A	-A

Exhibit F-9: Grade-Band 3-5 Social Studies Field-Test Classical Item Statistics

ITS Item ID	Grade	Item Position	Adjusted Biserial/ Polyserial	Average Score	Access Limitation	Omit	DIF	
							Female vs. Male	Black vs. White
1848	3-5	17	0.46	0.3946015	0.00	0.03	+A	-A
1849	3-5	18	0.51	0.4922880	0.00	0.02	+A	-A
1850	3-5	19	0.45	0.5760310	0.00	0.02	+A	-A
1851	3-5	20	0.47	0.3994845	0.00	0.03	+A	-A
1852	3-5	21	0.44	0.5386595	0.00	0.03	+A	+A
1853	3-5	22	0.47	0.5193300	0.00	0.02	+A	+A
1854	3-5	23	0.47	0.4922280	0.00	0.03	+A	-A
1855	3-5	33	0.58	0.8158190	0.00	0.01	-A	+A
1856	3-5	34	0.60	0.6007750	0.00	0.01	+A	+A
1857	3-5	35	0.54	0.7851605	0.00	0.00	+A	+A
1858	3-5	36	0.40	0.4150945	0.00	0.01	-A	-A
1859	3-5	37	0.67	0.7386235	0.00	0.01	+A	-A
1860	3-5	38	0.60	0.8324085	0.00	0.01	-A	+A
1861	3-5	39	0.56	0.7791990	0.00	0.01	+A	-A
1862	3-5	40	0.73	0.6422990	0.00	0.02	-A	-A
1875	3-5	56	0.34	0.5951090	0.00	0.00	+A	+A
1876	3-5	57	0.56	0.6115650	0.00	0.00	-A	+A
1877	3-5	58	0.68	0.6530615	0.00	0.00	-A	+A
1878	3-5	59	0.44	0.6163265	0.00	0.00	-A	-A
1879	3-5	60	0.56	0.7653060	0.00	0.00	+A	+A
1880	3-5	61	0.50	0.6069485	0.00	0.00	-A	-A
1881	3-5	62	0.36	0.6355585	0.00	0.00	-A	+A

Exhibit F-10: Grade-Band 6-8 Social Studies Field-Test Classical Item Statistics

ITS Item ID	Grade	Item Position	Adjusted Biserial/ Polyserial	Average Score	Access Limitation	Omit	DIF	
							Female vs. Male	Black vs. White
1863	6-8	17	0.31	0.5374150	0.00	0.02	-A	-A
1864	6-8	18	0.44	0.5477815	0.00	0.02	+A	-A
1866	6-8	19	0.45	0.3801370	0.00	0.03	+A	+A
1867	6-8	20	0.45	0.4879725	0.00	0.03	-A	-A
1868	6-8	21	0.32	0.5275860	0.00	0.03	-A	-A
1869	6-8	30	0.43	0.7176545	0.00	0.00	-A	+A
1870	6-8	31	0.35	0.5332930	0.00	0.01	-A	+A
1871	6-8	32	0.39	0.6652540	0.00	0.01	-A	+A
1872	6-8	33	0.40	0.5738500	0.00	0.01	+A	+A
1873	6-8	34	0.55	0.5643205	0.00	0.01	-A	+A
1874	6-8	35	0.51	0.5856620	0.00	0.02	+A	+A
1882	6-8	50	0.46	0.6393440	0.00	0.00	+A	-A
1883	6-8	51	0.47	0.8647540	0.00	0.00	+A	-A
1884	6-8	52	0.57	0.6004100	0.00	0.00	+A	+A
1885	6-8	53	0.66	0.6680385	0.00	0.01	+A	+A
1886	6-8	54	0.47	0.7417580	0.00	0.00	-A	-A
1887	6-8	55	0.60	0.7046705	0.00	0.01	+A	-A

Exhibit F-11: ELA Field-Test WINSTEPS Item Statistics

ITS ID	MEASURE	COUNT	SCORE	ERROR	IN.MSQ	IN.ZSTD	OUT.MS	OUT.ZSTD
1770	0.334479	608	590	0.058915	1.07	1.7	1.07	1.47
1771	-0.32315	620	837	0.054919	1.03	0.67	1.02	0.36
1767	0.535491	612	514	0.051779	1.15	3.91	1.18	3.61
1772	0.712547	604	451	0.054318	1.15	3.37	1.19	3.64
1768	0.704536	593	445	0.054348	1.02	0.59	1.02	0.39
1769	0.215963	599	626	0.051892	1.1	2.72	1.14	3.01
1778	0.308484	1271	1701	0.044141	1.04	1.32	1.06	1.51
1774	0.883872	1266	1385	0.041801	1.39	9.9	1.4	9.9
1775	0.344123	1271	1740	0.041347	1.24	6.64	1.28	5.66
1776	1.021226	1273	1311	0.040995	1.3	8.73	1.32	8.59
1777	0.030227	1275	1972	0.042756	1.1	2.54	1.12	1.64
1773	1.277972	1255	1130	0.039098	1.4	9.9	1.46	9.9
1706	0.905318	1044	1266	0.041389	0.98	-0.6	0.94	-1.22
1707	0.261596	1045	1532	0.049229	0.94	-1.64	0.92	-1.52
1708	0.65612	1044	1365	0.04532	1.28	7.32	1.32	6.67
1709	1.238001	1042	1065	0.042559	1.16	4.68	1.19	4.76

ITS ID	MEASURE	COUNT	SCORE	ERROR	IN.MSQ	IN.ZSTD	OUT.MS	OUT.ZSTD
1710	0.962584	1041	1220	0.042053	1.02	0.49	1.02	0.45
1711	0.122924	1038	1658	0.049132	1.12	2.43	1.18	2.26
1784	-0.179269	474	636	0.063012	1.09	1.76	1.1	1.31
1779	0.637346	459	404	0.062245	1.11	2.42	1.15	2.31
1780	0.419539	467	464	0.059357	1.07	1.56	1.09	1.49
1782	0.80656	458	380	0.070256	1.21	3.78	1.24	3.9
1783	-0.0968	460	605	0.062039	1.1	1.86	1.11	1.41
1712	1.194464	1157	1272	0.041446	1.07	1.95	1.16	3.14
1713	0.385962	1153	1638	0.048465	0.97	-0.75	1.03	0.64
1714	1.328646	1152	1190	0.040896	1.04	1.21	1	0
1715	0.167883	1157	1749	0.049704	0.85	-3.97	0.8	-3.64
1717	0.524067	1146	1632	0.044505	1.17	4.36	1.25	3.61
1716	1.424661	1148	1146	0.046754	1.65	9.9	1.73	9.9
1830	1.087177	1057	1245	0.048212	1.36	8.94	1.44	9.61
1831	1.157546	1055	1243	0.042205	0.99	-0.26	0.95	-0.83
1832	1.353906	1058	1133	0.043114	0.94	-1.78	0.91	-1.96
1833	1.899805	1052	888	0.048599	1.5	9.9	1.55	9.9
1834	0.827829	1051	1339	0.05008	1.28	6.84	1.3	6.42
1835	0.793646	1045	1370	0.048212	1.18	4.57	1.18	3.53
1804	0.227856	128	146	0.135881	1.15	1.46	1.14	1.35
1806	0.571162	129	126	0.114562	1.16	1.78	1.16	1.28
1813	-0.252742	125	180	0.130552	1.03	0.32	1.03	0.26
1808	0.411649	128	136	0.132841	0.86	-1.48	0.84	-1.64
1810	0.341361	127	142	0.118763	0.99	-0.04	0.96	-0.33
1766	-0.025133	306	507	0.104058	0.75	-2.82	0.65	-2.4
1823	-0.174596	306	511	0.109449	0.79	-2.3	0.65	-2.61
1825	1.18986	305	358	0.081508	0.97	-0.42	0.93	-0.75
1826	0.422899	308	462	0.091572	1.02	0.29	1.02	0.24
1827	0.476398	306	461	0.088824	1.18	2.21	1.12	0.82
1803	0.66418	273	404	0.094307	0.97	-0.4	0.94	-0.41
1800	1.022859	272	353	0.09189	1.09	1.26	1.12	1.25
1801	2.167744	273	212	0.091948	1.34	4.4	1.47	5.04
1802	1.100009	272	351	0.087043	1.4	4.98	1.44	3.73
1799	1.200782	271	331	0.090535	0.95	-0.64	0.9	-1.18
1798	1.650185	270	274	0.085588	1.45	5.95	1.55	5.63

Exhibit F-12: Math Field-Test WINSTEPS Item Statistics

ITS ID	MEASURE	COUNT	SCORE	ERROR	IN.MSQ	IN.ZSTD	OUT.MS	OUT.ZSTD
1634	-0.883584	696	1491	0.046202	0.81	-3.77	0.83	-2.74
1635	-0.185085	646	358	0.083066	1	0.15	1.02	0.5
1636	-0.823715	655	450	0.088216	1.08	2.06	1.1	2.08
1637	0.311022	643	548	0.051525	1.17	4.21	1.23	4.07
1638	-0.277952	639	739	0.058593	1.04	0.94	1.04	0.95
1639	-0.146125	639	718	0.051488	1.17	4.06	1.2	3.65
1646	0.596405	1261	1439	0.043642	1.2	5.76	1.2	4.94
1647	0.388946	1268	1617	0.03909	1.23	6.69	1.28	4.46
1648	0.570246	1258	1455	0.042851	1.13	3.91	1.13	3.31
1649	0.797173	1269	1339	0.039448	1.13	3.87	1.14	3.08
1650	0.379185	1261	1635	0.038237	1.17	5.02	1.07	1.07
1651	0.402885	1254	1531	0.044411	1.07	2.07	1.06	1.46
1640	1.925091	1043	352	0.070774	1.28	8.39	1.41	9.32
1641	0.069626	1037	748	0.072797	1.03	0.93	1.02	0.28
1642	0.133915	1038	1580	0.04645	0.97	-0.62	1.07	0.9
1643	0.754227	1039	1199	0.050023	1.13	3.41	1.13	3.11
1644	0.527194	1039	1361	0.044093	1.27	7.25	1.48	7.53
1645	0.334583	1036	1495	0.043434	0.92	-2.28	0.82	-2.35
1730	0.341665	462	431	0.066482	1.14	2.7	1.19	3.29
1731	-0.572972	451	302	0.105406	1	0.03	0.99	-0.08
1732	0.531326	457	381	0.065211	1.19	3.83	1.29	4.6
1733	-0.091985	450	256	0.100187	0.99	-0.23	0.98	-0.59
1734	0.717084	446	315	0.06307	1.11	2.24	1.13	1.67
1735	0.419003	460	416	0.071284	1.12	2.3	1.13	2.39
1718	1.29585	1157	555	0.063821	1.09	4.11	1.12	3.62
1719	1.081307	1154	606	0.063829	1.05	2.19	1.05	1.42
1722	1.20594	1156	1157	0.046585	1.41	9.9	1.45	9.9
1723	1.201673	1154	1149	0.04138	1.26	7.44	1.33	7.02
1724	0.84908	1031	1275	0.044719	1.39	9.9	1.53	9.59
1726	0.981085	1026	1212	0.043453	1.42	9.9	1.52	9.45
1725	0.903288	1026	1216	0.048512	1.14	3.7	1.15	3.46
1727	1.379458	1025	997	0.042896	1.14	4.1	1.18	3.84
1728	0.944881	1020	1227	0.043533	1.32	8.61	1.34	6.26
1729	1.180614	1016	1093	0.043351	1.43	9.9	1.5	9.71
1759	-0.51838	131	287	0.108175	0.72	-2.02	0.72	-1.5
1761	0.430074	127	121	0.112699	1.18	2.22	1.17	1.22
1762	0.389635	125	122	0.111107	1.06	0.81	1.12	0.93
1763	0.474855	127	115	0.114195	1.09	1.06	1.13	1.05
1764	-0.102951	126	159	0.121264	1.15	1.5	1.14	1.07

ITS ID	MEASURE	COUNT	SCORE	ERROR	IN.MSQ	IN.ZSTD	OUT.MS	OUT.ZSTD
1752	0.950203	301	320	0.092138	1.19	2.6	1.18	2.39
1753	0.936984	300	324	0.081747	1.3	4.66	1.43	5.19
1754	0.30614	302	426	0.083851	1.02	0.31	0.97	-0.27
1755	0.846812	303	341	0.079584	1.02	0.31	1.04	0.5
1756	0.839788	301	344	0.075491	1.1	1.8	1.18	1.91
1757	0.933141	300	326	0.078591	1.39	5.98	1.5	5.65
1758	1.222495	301	276	0.074995	1.5	7.49	1.68	6.57
1742	0.672935	265	326	0.093417	1.01	0.1	0.99	-0.12
1743	0.507968	265	370	0.083087	1.13	1.79	1.2	1.52
1744	1.331809	263	237	0.078583	1.19	2.92	1.23	2.3
1745	0.27891	263	390	0.091208	0.94	-0.71	1.31	2.23
1746	-0.177985	265	448	0.106273	0.87	-1.1	0.77	-1.24

Exhibit F-13: Science Field-Test WINSTEPS Item Statistics

ITS ID	MEASURE	COUNT	SCORE	ERROR	IN.MSQ	IN.ZSTD	OUT.MS	OUT.ZSTD
894	-1.00231	518	1234	0.05961	0.81	-2.85	0.8	-2.15
1023	-0.808851	487	360	0.108112	0.91	-1.74	0.83	-2.42
895	-0.558666	486	337	0.103286	0.93	-1.54	0.89	-1.79
896	-0.085402	485	289	0.097254	0.97	-0.85	0.94	-1.33
898	0.087432	470	531	0.068862	1.05	0.92	1.05	0.88
965	-0.671091	481	346	0.105869	0.95	-0.94	0.9	-1.57
966	-0.330087	485	674	0.063604	0.92	-1.53	0.9	-1.24
967	-0.370694	482	318	0.100678	1.01	0.33	1.01	0.27
968	0.174069	477	517	0.069515	1.01	0.16	0.99	-0.16
969	-0.194152	471	297	0.099613	0.99	-0.21	0.96	-0.77
1663	-0.402094	482	656	0.069492	0.97	-0.49	0.92	-1.3
1658	-0.117989	476	583	0.068948	1.02	0.48	1.01	0.14
1659	-0.037244	477	584	0.062156	1.01	0.14	0.98	-0.28
1660	-0.370205	482	643	0.070172	0.93	-1.36	0.91	-1.59
1661	0.297602	478	498	0.058717	1.04	0.88	1.06	1.05
1662	0.127829	472	544	0.059779	1.11	2.42	1.15	2.46
1669	0.016404	876	625	0.078288	1.01	0.19	0.99	-0.12
1665	-0.256861	876	1438	0.056226	0.86	-2.66	0.66	-4.09
1828	0.235552	877	1209	0.051089	0.83	-4.43	0.76	-5.02
1666	-0.340497	875	1390	0.058965	0.83	-3.55	0.72	-4.6
1829	1.050845	875	862	0.045292	1.01	0.35	1.04	0.88
1668	0.320743	871	1154	0.051952	1.12	2.9	1.11	2.19
1670	0.325321	657	986	0.058088	0.99	-0.11	0.93	-0.83
1673	0.212489	653	966	0.062841	0.88	-2.51	0.83	-2.75
1675	0.077936	654	1024	0.063903	1.09	1.62	1.13	1.59

ITS ID	MEASURE	COUNT	SCORE	ERROR	IN.MSQ	IN.ZSTD	OUT.MS	OUT.ZSTD
1672	1.054259	651	725	0.052597	1.03	0.85	1.05	1.11
1671	0.61009	652	893	0.053869	1.31	6.44	1.46	5.92
1674	1.508353	653	584	0.059745	1.33	6.78	1.33	6.71
1676	-0.443915	378	506	0.07822	1.01	0.24	0.97	-0.4
1677	-0.183139	382	488	0.06764	1.25	4.3	1.29	2.99
1678	0.329947	376	365	0.06761	1.01	0.18	1	0.07
1679	0.267311	380	381	0.076899	1.21	3.46	1.22	3.26
1680	0.531874	374	336	0.080938	1.27	4.17	1.3	4.21
1681	0.660469	366	290	0.070306	1.16	2.8	1.24	2.74
957	-0.927886	886	743	0.097542	0.96	-0.6	1.09	0.83
959	0.330648	874	1186	0.049046	0.81	-4.91	0.67	-4.55
958	0.07766	877	605	0.079665	0.86	-4.28	0.77	-4.21
960	-0.104447	865	1313	0.055578	0.96	-0.76	0.82	-2.25
963	0.001526	875	1315	0.052459	0.98	-0.31	0.83	-1.87
1682	0.119474	795	1173	0.056826	0.9	-2.18	0.79	-3.16
1683	0.263561	798	1153	0.053353	0.88	-2.72	0.8	-2.82
1685	1.24706	792	781	0.059098	1.38	7.91	1.4	8.28
1686	1.166011	792	798	0.050671	1.05	1.27	1.04	0.88
1696	0.437309	693	1038	0.054217	0.73	-6.07	0.56	-4.67
1694	0.264301	692	966	0.066221	1.1	1.99	1.07	1.28
1695	0.004257	693	1052	0.067079	0.94	-1.23	0.9	-1.56
1697	0.752478	689	900	0.054047	0.96	-0.85	0.96	-0.58
1698	1.037044	691	797	0.05407	1.22	4.9	1.28	4.8
1699	0.630057	687	912	0.058343	1.05	1.01	1.03	0.6
1003	0.035439	649	501	0.098179	0.86	-2.89	0.73	-3.7
1004	1.937034	644	487	0.059063	1.8	9.9	2.18	9.9
1005	1.390849	646	654	0.055928	1.09	1.93	1.09	1.75
1006	1.238025	648	350	0.085049	0.99	-0.21	0.97	-0.83

Exhibit F-14: Social Studies Field-Test WINSTEPS Item Statistics

ITS ID	MEASURE	COUNT	SCORE	ERROR	IN.MSQ	IN.ZSTD	OUT.MS	OUT.ZSTD
1848	0.306977	365	310	0.067529	1.07	1.39	1.2	2.54
1849	-0.043678	373	387	0.078378	0.99	-0.1	0.98	-0.25
1850	-0.337612	367	453	0.069397	1.12	2.1	1.25	3.14
1851	0.309197	365	310	0.068016	1.09	1.77	1.09	1.26
1852	-0.19944	365	424	0.067645	1.15	2.81	1.19	2.5
1853	-0.138471	367	408	0.068495	1.08	1.59	1.08	1.13
1854	-0.052665	362	383	0.075468	1.06	1.17	1.05	0.82
1855	-0.380573	898	1481	0.05686	0.98	-0.4	0.85	-1.54
1856	0.600211	898	1087	0.045748	1.05	1.18	1.03	0.46
1857	-0.203827	902	1424	0.053615	1.04	0.75	1	-0.01
1858	1.333623	896	748	0.048288	1.27	6.45	1.44	8.21
1859	-0.056708	896	1335	0.053029	0.88	-2.79	0.77	-3.6
1860	-0.422872	898	1504	0.057786	0.96	-0.69	0.84	-1.5
1861	-0.184642	896	1407	0.053743	1.02	0.45	0.92	-0.9
1862	0.467122	890	1153	0.045798	0.86	-3.89	0.75	-3.85
1875	0.818833	734	878	0.052902	1.25	5.76	1.24	4.47
1876	0.813639	733	899	0.049509	1.04	1.04	1.01	0.2
1877	0.675497	733	960	0.049596	0.87	-3.29	0.81	-2.8
1878	0.741278	735	907	0.053195	1.11	2.59	1.15	2.77
1879	0.196252	735	1127	0.055836	0.95	-0.93	0.88	-1.34
1880	0.743805	733	892	0.055401	1.02	0.39	1.02	0.52
1881	0.566918	733	934	0.058292	1.13	2.97	1.13	2.72
1863	-0.107413	271	316	0.080794	1.3	4.45	1.35	4.16
1864	-0.240767	272	321	0.093399	1.06	0.88	1.05	0.77
1866	0.473749	267	222	0.081499	1.09	1.39	1.14	1.71
1867	0.058773	270	284	0.092885	1.04	0.63	1.05	0.71
1868	-0.09048	265	306	0.082165	1.27	4.02	1.33	3.92
1869	0.176741	827	1188	0.056774	1.09	2.07	1.2	3.14
1870	1.056569	820	881	0.051634	1.36	8.27	1.39	7.5
1871	0.536436	817	1101	0.050189	1.29	6.47	1.47	5.91
1872	0.868029	820	949	0.052293	1.25	5.99	1.33	6.37
1873	0.949757	817	930	0.047488	1.11	2.81	1.1	1.63
1874	0.815222	813	964	0.052427	1.1	2.44	1.07	1.51
1882	0.683102	732	468	0.082084	0.98	-0.84	0.94	-1.17
1883	-0.20302	731	1267	0.067332	0.96	-0.46	0.85	-1.04
1884	0.912294	732	879	0.054033	0.95	-1.18	0.94	-1.17
1885	0.711752	727	974	0.051146	0.82	-4.56	0.72	-4.09
1886	0.373483	729	1081	0.055905	1.05	1.03	0.97	-0.39
1887	0.267644	725	1027	0.063284	0.86	-3.21	0.82	-3.44

Appendix G: Marginal Reliability by Grade-Band, Subject, Starting Task, Gender, and Ethnic Group

Exhibit G-1: Marginal Reliability by Starting Task and Grade-Band for ELA

Initial Task	Statistic	Elementary	Middle	High
1	N	367	270	70
	$\bar{\sigma}_{e^*}$	14.51	13.97	18.45
	Reliability	0.90	0.91	0.92
3	N	306	244	70
	$\bar{\sigma}_{e^*}$	10.38	10.78	14.71
	Reliability	0.87	0.88	0.86
7	N	796	796	209
	$\bar{\sigma}_{e^*}$	15.57	25.81	16.56
	Reliability	0.81	0.79	0.84

Exhibit G-2: Marginal Reliability by Starting Task and Grade-Band for Mathematics

Initial Task	Statistic	Elementary	Middle	High
1	N	354	241	67
	$\bar{\sigma}_{e^*}$	14.95	16.25	18.37
	Reliability	0.89	0.86	0.91
3	N	354	277	72
	$\bar{\sigma}_{e^*}$	11.64	12.09	14.53
	Reliability	0.83	0.80	0.86
7	N	755	778	203
	$\bar{\sigma}_{e^*}$	20.06	20.70	16.31
	Reliability	0.79	0.82	0.85

Exhibit G-3: Marginal Reliability by Starting Task and Grade-Band for Science/Biology

Initial Task	Statistic	Elementary	Middle	High
1	N	293	214	206
	$\bar{\sigma}_{e^*}$	17.29	17.16	26.96
	Reliability	0.89	0.90	0.84
3	N	225	196	64
	$\bar{\sigma}_{e^*}$	15.42	13.54	29.65
	Reliability	0.78	0.85	0.81
7/6*	N	501	494	50
	$\bar{\sigma}_{e^*}$	18.87	28.60	38.04
	Reliability	0.75	0.74	0.74

**

Initial task for elementary school and middle school science is Task 7; for high school, it is Task 6.

Exhibit G-4: Marginal Reliability by Starting Task and Grade-Band for Social Studies

Initial Task	Statistic	Elementary	Middle	High
1	N	216	144	X
	$\bar{\sigma}_{e^*}$	16.95	18.36	X
	Reliability	0.89	0.88	X
3	N	186	160	X
	$\bar{\sigma}_{e^*}$	13.75	16.05	X
	Reliability	0.82	0.82	X
7	N	603	588	X
	$\bar{\sigma}_{e^*}$	17.60	20.08	X
	Reliability	0.80	0.79	X

Exhibit G-5: Marginal Reliability by Subject, Gender, and Grade-Band

Subject	Gender	Elementary	Middle	High	Overall
ELA	Female	0.935	0.909	0.938	0.925
	Male	0.930	0.892	0.922	0.911
Math	Female	0.927	0.921	0.925	0.926
	Male	0.916	0.909	0.918	0.914
Science	Female	0.911	0.899		0.905
	Male	0.894	0.882		0.887
Biology	Female			0.849	0.849
	Male			0.872	0.872
Social Studies	Female	0.926	0.917		0.922
	Male	0.924	0.908		0.916

Exhibit G-6: Marginal Reliability by Subject, Major Ethnic Group, and Grade-Band

Subject	Ethnicity	Elementary	Middle	High	Overall
ELA	African American	0.930	0.891	0.921	0.910
	White	0.933	0.901	0.935	0.920
Math	African American	0.916	0.914	0.913	0.915
	White	0.923	0.909	0.930	0.919
Science	African American	0.896	0.892		0.894
	White	0.902	0.875		0.889
Biology	African American			0.875	0.875
	White			0.856	0.856
Social Studies	African American	0.924	0.910		0.917
	White	0.925	0.911		0.919

Appendix H: Score Report Sample

INDIVIDUAL STUDENT REPORT

Prepared Especially for the Family of
Kyree Adams

Date of Birth: 9/20/2000
Student ID: 587412589457
School District: Calvert
School: Alfonso Elementary School

Spring 2011



The South Carolina Alternate Assessment (SC-Alt)

Kyree participated in the South Carolina Alternate Assessment (SC-Alt) during the spring of 2011.

She took the elementary school form of the test, which is based on academic standards from grades 3 to 5. This report is designed to provide you with information on your child's performance on this assessment.

The SC-Alt is a test designed for students with significant cognitive disabilities who participate in a school curriculum that includes academic and functional skill instruction. The alternate assessment only tests students' achievement in English language arts (ELA), mathematics, science, and social studies. Individualized Education Program (IEP) reports and other methods provide parents with information on how students are progressing in the other areas.

What is the SC-Alt?

- The SC-Alt includes performance tasks in each subject area. Students may complete the tasks by using their usual method of communication. This may include pointing or gazing at answer choices; selecting objects, pictures, or picture symbols that represent an answer choice; or reading letters, words or sentences to complete the task.
- The tasks are linked to the state academic content standards in four areas: English language arts (ELA), mathematics, science, and social studies.
- Students are assigned a test form based on their age. Students ages 8-10 are assigned to the elementary school form; students ages 11-13 are assigned to the middle school form; and students age 15 take the high school form.

How are scores reported and used?

- Four achievement levels (Level 1, Level 2, Level 3, and Level 4) have been established for the SC-Alt. Achievement levels describe how students are doing in relation to the state academic standards. Your child's performance is also reported as a scale score that allows parents to monitor growth from year to year.
- Level 2 is the achievement level reported as "Met" on the District Report Card for state accountability. Levels 3 and 4 are the achievement levels reported as "proficient" in ELA and mathematics for schools' and districts' Adequate Yearly Progress (AYP) report for federal accountability.

Where can I get more information about SC-Alt and my child's performance?

- You can contact your child's teacher or school for more information.
- You can view examples of tasks, information about expectations at each achievement level, and scale score tables on the South Carolina Department of Education website at <http://ed.sc.gov/agency/Accountability/Assessment/SouthCarolinaAlternateAssessment.html>.



The South Carolina Department of Education

Spring 2011
Kyree Adams

The SC-Alt



Mathematics

Kyree scored at **Level 3** with a scale score of **480** in mathematics.

Students who score at Level 3 should be able to:

- add and subtract simple numbers;
- count and compare objects in a group;
- compare objects by color, size, or shape;
- identify three-dimensional shapes;
- read information in a graph.

Your
Child's
Level

- | | |
|---|--|
| 4 | Students performing at Level 4 demonstrate and apply academic skills and competencies in mathematics. |
| 3 | Students performing at Level 3 demonstrate increasing academic skills and competencies in mathematics. |
| 2 | Students performing at Level 2 demonstrate foundational academic skills and competencies in mathematics. |
| 1 | Students performing at Level 1 may demonstrate emerging academic skills and competencies in mathematics. |

How you can support Kyree's learning

- Help your child add and subtract during everyday activities. For example, show her five pennies or other objects and give her one more. Then, ask her to tell you how many there are altogether.
- Play games with your child. Use games that require matching numbers (dots), such as dominoes.
- Put three objects in one pile and two objects in another pile. Ask your child which pile has more objects.
- Describe everyday household objects by shapes. For example, a can is a cylinder; a box is a rectangular prism; and a ball is a sphere.



English Language Arts

Kyree scored at **Level 4** with a scale score of **495** in ELA.

Students who score at Level 4 should be able to:

- identify the main idea and make predictions about what will happen next in a story;
- write a simple story;
- follow two-step directions;
- take turns appropriately during conversations.

Your
Child's
Level

- | | |
|---|---|
| 4 | Students performing at Level 4 demonstrate and apply academic skills and competencies in reading, writing, and research. |
| 3 | Students performing at Level 3 demonstrate increasing academic skills and competencies in reading, writing, and research. |
| 2 | Students performing at Level 2 demonstrate foundational academic skills and competencies in reading, writing, and research. |
| 1 | Students performing at Level 1 may demonstrate emerging academic skills and competencies in reading, writing, and research. |

How you can support Kyree's learning

- Encourage your child to read passages from a variety of materials (books, magazines, newspapers).
- Read a story with your child and talk to her about specific characters and the order of events in the story.
- Assist your child with writing about an event or activity using her typical method of communication.
- Encourage your child to begin conversations with family members or friends by using her typical method of communication.

Spring 2011
Kyree Adams

Science

Kyree scored at **Level 1** with a scale score of **270** in science.

Students who score at Level 1 should be able to:

- use senses to observe the outcome of a simple scientific investigation;
- sequence growth patterns;
- observe and record daily weather conditions;
- recognize the sun and moon and relate them to day and night;
- recognize that objects move when force is applied.

4	Students performing at Level 4 demonstrate and apply academic skills and competencies in science.
3	Students performing at Level 3 demonstrate increasing academic skills and competencies in science.
2	Students performing at Level 2 demonstrate foundational academic skills and competencies in science.
1	Students performing at Level 1 may demonstrate emerging academic skills and competencies in science.

Your Child's Level →

How you can support Kyree's learning

- Help your child plant a seed and watch it grow.
- Explain and help your child observe different weather conditions.
- Point out the moon in the sky at night.
- Show your child two balls. Roll one across a table and leave one still. Point out which ball is in motion and which is not.



Social Studies

Kyree scored at **Level 3** with a scale score of **500** in Social Studies.

Students who score at Level 3 should be able to:

- understand the concept of past and present;
- demonstrate respect for people of authority;
- identify major symbols of the United States (the flag, bald eagle);
- recognize that when we work we earn money to buy things;
- identify features on a map of South Carolina (river, mountain, ocean);
- match historical figures such as Abraham Lincoln, Thomas Jefferson, etc., to their accomplishments.

4	Students performing at Level 4 demonstrate and apply academic skills and competencies in social studies.
3	Students performing at Level 3 demonstrate increasing academic skills and competencies in social studies.
2	Students performing at Level 2 demonstrate foundational academic skills and competencies in social studies.
1	Students performing at Level 1 may demonstrate emerging academic skills and competencies in social studies.

Your Child's Level →

How you can support Kyree's learning

- Play a "Now or Long Ago" game (e.g., do we ride in a stagecoach now or long ago?).
- Find and identify items in the community such as the American flag.
- Talk about national holidays (Fourth of July and what it means) and celebrate them with your child.
- Look at a map of South Carolina with your child and find what is near the mountains and near the ocean.
- Let your child earn an allowance by doing chores or helping out by following rules. (A job can be as simple as not throwing a tantrum or allowing you to brush her teeth without fussing.)

Spring 2011
Kelly Adams

The SC-Alt



The following areas are tested in Mathematics:

Number and Operations

- whole numbers
- fractions
- addition and subtraction
- multiplication and division

Algebra

- patterns and their relationships

Geometry

- attributes of objects such as shape, size, color
- identification of two- and three-dimensional shapes

Measurement

- money
- length, liquid, volume, and mass and weight
- time
- equivalences

Data Analysis and Probability

- data collection and representation
- data analysis
- probability



The following areas are tested in English Language Arts:

Reading

- reading
 - comprehending a variety of texts (such as fiction, nonfiction, poetry, and drama)
- Note: Reading materials may include objects, pictures or photographs, picture symbols, letters, and words.*

Writing

- developing written communications (notes, stories) using the student's typical method of communication

Research

- accessing and using information from a variety of sources
 - communicating their own ideas and ideas of others*
- *Students' typical method of communication, verbal or nonverbal, may be facilitated by using objects, pictures or photographs, picture symbols, letters and words, voice output devices, or assistive technology.*



The following areas are tested in Science:

Scientific Inquiry involves studying scientific processes and skills such as:

- observing
- classifying
- predicting what will happen in a simple scientific experiment

Life Science

- basic needs of plants and animals
- their structures and habitats

Earth Science

- weather
- objects in the sky (sun and moon)
- earth materials (rocks and soil)

Physical Science

- characteristics of objects
- the effect of force on the motion of objects
- light, heat, and electricity



The following areas are tested in Social Studies:

Social Studies Literacy Elements are concepts required for understanding this subject such as:

- distinguishing between past, present, and future
- demonstrating responsible citizenship within the school community, the local community, and national communities
- creating and using timelines
- understanding the relationship between people and the land

Academic Standards include concepts related to specific historical time frames:

- history
- geography
- political science/government
- economics

Appendix I: Student Performance by Demographics, Grade-Band, and Subject Area

Exhibit I-1: Performance by Grade-Band and Demographics—ELA¹²

	Elementary School (ES)					Middle School (MS)					High School (HS)				
	N	Scale Score		Ach. Level		N	Scale Score		Ach. Level		N	Scale Score		Ach. Level	
		Mean	SD	< 3	>= 3		Mean	SD	< 3	>= 3		Mean	SD	< 3	>= 3
STUDENT'S AGE															
8	417	486	55	134	281
9	508	503	56	101	406
10	446	501	65	111	329	1	—	—	—	—
11	425	517	71	108	315
12	430	515	68	110	318
13	355	512	76	96	258
15	312	513	68	93	217
16	15	533	80	4	11
STUDENT'S ETHNICITY															
Unknown	1	—	—	—	—
Asian	22	485	42	8	14	11	485	49	2	9	5	520	31	.	5
African American	726	505	60	155	566	684	522	75	161	522	180	516	71	55	124
Hispanic	89	483	60	32	57	50	511	62	14	36	5	488	81	2	3
American Indian	4	—	—	—	—	3	—	—	—	—	1	494	.	.	1
Other	41	481	59	16	24	30	489	59	13	17	7	566	80	2	5
Hawaiian/Pacific Islander	2	—	—	—	—	1	—	—	—	—	1	533	.	.	1
White	611	494	58	161	446	552	514	69	143	405	158	518	67	46	111
STUDENT'S GENDER															
Unknown	1
Female	479	495	62	132	346	432	513	74	107	321	115	521	72	31	84
Male	1016	500	58	241	766	899	519	71	227	671	242	516	68	74	166
ESL (LANGUAGE)															
Unknown	1	—	—	—	—	2	—	—	—	—
Pre-functional	60	485	64	20	40	35	491	72	13	22	4	—	—	—	—
Beginner	3	—	—	—	—	2	—	—	—	—
Intermediate
Advanced
Full English Proficient
Title III First Year Exited
Title III Second+ Year Exited

¹² Note: Data marked '—' are suppressed because the subgroup contains fewer than 10 students.

	Elementary School (ES)					Middle School (MS)					High School (HS)				
	N	Scale Score		Ach. Level		N	Scale Score		Ach. Level		N	Scale Score		Ach. Level	
		Mean	SD	< 3	>= 3		Mean	SD	< 3	>= 3		Mean	SD	< 3	>= 3
English Speaker I	1	–	–	–	–	2	–	–	–	–	1	–	–	–	–
English Speaker II	1430	499	59	352	1068	1289	518	72	320	964	352	518	69	104	246
Pre-functional Waiver	1	–	–	–	–
Beginner Waiver	1	–	–	–	–
Intermediate Waiver
Advanced Waiver
ELIGIBLE FOR FREE OR REDUCED-PRICE LUNCH															
Full-Pay Meals	395	479	57	133	256	388	504	66	117	267	112	499	68	40	71
Free Meals	997	506	57	216	777	845	524	75	186	658	224	526	69	57	166
Reduced Meals	103	496	64	24	79	98	504	54	31	67	21	519	66	8	13
Unknown	1	–	–	–	–
IEP Disability Codes (Multiple Codes per Student)															
Severe Mental Disability	136	408	63	121	14	136	427	65	110	24	31	410	81	25	6
Moderate Mental Disability	333	482	43	90	242	404	501	48	101	303	133	512	50	45	87
Mild Mental Disability	389	534	39	17	370	365	564	56	13	351	98	564	40	3	94
Autism	400	492	45	111	286	302	504	59	100	201	70	498	69	29	41
Deaf/Blindness															
Emotional Disability	17	553	48	1	14	10	556	59	1	9	1	–	–	–	–
Hearing Impaired	21	471	75	8	13	21	506	60	5	16	11	531	33	1	10
Learning Disability	44	559	39	1	43	27	635	62	.	27	3	–	–	–	–
Multiple-Disability	1	–	–	–	–	2	–	–	–	–	2	–	–	–	–
Other Health Impaired	79	502	62	19	59	63	544	65	7	56	14	540	51	2	12
Orthopedically Impaired	61	485	45	25	36	49	497	77	17	31	14	526	45	2	12
Speech or Language Impaired	1090	500	51	238	847	665	513	60	169	492	119	522	53	31	87
Traumatic Brain Injury	11	492	52	5	6	8	–	–	–	–	4	–	–	–	–
Visually Impaired	53	443	66	35	17	49	475	54	27	21	11	513	82	4	7
TOTAL	1496	498	59	373	1112	1331	517	72	334	992	357	518	69	105	250

Exhibit I-2: Performance by Grade-Band and Demographics—Mathematics¹³

	Elementary School (ES)					Middle School (MS)					High School (HS)				
	N	Scale Score		Ach. Level		N	Scale Score		Ach. Level		N	Scale Score		Ach. Level	
		Mean	SD	< 3	>= 3		Mean	SD	< 3	>= 3		Mean	SD	< 3	>= 3
STUDENT'S AGE															
8	417	487	62	156	259
9	508	506	63	138	369
10	446	503	73	135	305	1	–	–	–	–
11	425	516	63	131	290
12	430	513	67	132	295
13	355	510	74	110	243
15	312	506	67	116	190
16	15	509	74	5	10
STUDENT'S ETHNICITY															
Unknown	1	–	–	–	–
Asian	22	495	65	8	14	11	497	62	3	8	5	–	–	–	–
African American	726	507	67	192	529	684	519	70	191	492	180	510	70	61	116
Hispanic	89	487	67	34	55	50	504	61	18	32	5	–	–	–	–
American Indian	4	–	–	–	–	3	–	–	–	–	1	–	–	–	–
Other	41	485	65	16	24	30	493	60	14	16	7	537	59	2	5
Hawaiian/Pacific Islander	2	–	–	–	–	1	–	–	–	–	1	–	–	–	–
White	611	494	65	206	401	552	514	68	174	368	158	507	62	60	94
STUDENT'S GENDER															
Unknown	1
Female	479	495	67	157	321	432	508	75	134	292	115	504	69	38	75
Male	1016	502	66	300	707	899	518	65	267	627	242	510	65	90	147
ESL (LANGUAGE)															
Unknown	1	–	–	–	–	2	–	–	–	–
Pre-functional	60	489	68	20	40	35	487	69	16	19	4	–	–	–	–
Beginner	3	–	–	–	–	2	–	–	–	–
Intermediate
Advanced
Full English Proficient
Title III First Year Exited
Title III Second+ Year Exited

¹³ Note: Data marked '–' are suppressed because the subgroup contains fewer than 10 students.

	Elementary School (ES)					Middle School (MS)					High School (HS)				
	N	Scale Score		Ach. Level		N	Scale Score		Ach. Level		N	Scale Score		Ach. Level	
		Mean	SD	< 3	>= 3		Mean	SD	< 3	>= 3		Mean	SD	< 3	>= 3
English Speaker I	1	—	—	—	—	2	—	—	—	—	1	—	—	—	—
English Speaker II	1430	500	66	436	984	1289	516	69	384	894	352	509	66	127	218
Pre-functional Waiver	1	—	—	—	—
Beginner Waiver	1	—	—	—	—
Intermediate Waiver
Advanced Waiver
ELIGIBLE FOR FREE OR REDUCED-PRICE LUNCH															
Full-Pay Meals	395	480	64	160	229	388	504	63	142	242	112	487	66	53	56
Free Meals	997	508	65	261	732	845	521	72	231	607	224	519	64	64	156
Reduced Meals	103	497	76	36	67	98	508	54	28	70	21	511	56	11	10
Unknown	1	—	—	—	—
IEP Disability Codes (Multiple Codes per Student)															
Severe Mental Disability	136	400	67	122	13	136	419	69	120	14	31	407	87	26	5
Moderate Mental Disability	333	481	46	129	203	404	500	45	143	260	133	503	43	64	66
Mild Mental Disability	389	540	50	27	360	365	559	52	16	346	98	552	49	5	90
Autism	400	494	49	144	253	302	510	55	104	197	70	494	60	29	40
Deaf/Blindness
Emotional Disability	17	539	49	1	14	10	546	43	1	9	1	—	—	—	—
Hearing Impaired	21	470	74	8	13	21	508	49	7	13	11	531	35	1	10
Learning Disability	44	567	54	2	42	27	609	56	.	27	3	—	—	—	—
Multiple-Disability	1	—	—	—	—	2	—	—	—	—	2	—	—	—	—
Other Health Impaired	79	505	77	21	57	63	542	63	8	52	14	522	42	4	10
Orthopedically Impaired	61	480	55	26	35	49	490	69	23	25	14	500	40	6	8
Speech or Language Impaired	1090	502	57	309	776	665	514	57	210	446	119	515	51	36	80
Traumatic Brain Injury	11	493	51	4	7	8	—	—	—	—	4	—	—	—	—
Visually Impaired	53	436	81	37	15	49	472	59	31	17	11	489	74	6	5
TOTAL	1496	500	66	457	1028	1331	515	69	401	919	357	508	66	128	222

Exhibit I-3: Performance by Grade-Band and Demographics—Science/Biology¹⁴

	Elementary School (ES)					Middle School (MS)					High School (HS)				
	N	Scale Score		Ach. Level		N	Scale Score		Ach. Level		N	Scale Score		Ach. Level	
		Mean	SD	< 3	>= 3		Mean	SD	< 3	>= 3		Mean	SD	< 3	>= 3
STUDENT'S AGE															
8	417	491	60	57	147
10	446	507	71	47	193	1	–	–	–	–
11	425	520	72	72	140
12	430	513	73	147	277
13	355	516	77	61	138
15	312	504	99	120	185
16	15	549	120	4	11
STUDENT'S ETHNICITY															
Unknown	1	–	–	–	–
Asian	22	492	37	3	10	11	496	8	1	5	5	–	–	–	–
African American	726	511	64	93	403	684	519	77	159	325	180	509	103	63	115
Hispanic	89	491	64	19	44	50	507	59	9	20	5	–	–	–	–
American Indian	4	–	–	–	–	3	–	–	–	–	1	–	–	–	–
Other	41	496	44	5	24	30	478	61	12	10	7	–	–	–	–
Hawaiian/Pacific Islander	2	–	–	–	–	1	–	–	–	–	1	–	–	–	–
White	611	501	60	83	344	552	521	72	116	254	158	509	99	63	90
STUDENT'S GENDER															
Unknown	1	–	–	–	–
Female	479	497	66	82	257	432	509	79	104	185	115	510	101	41	71
Male	1016	508	60	122	570	899	522	72	195	430	242	509	101	91	147
ESL (LANGUAGE)															
Unknown	1	–	–	–	–	2	–	–	–	–
Pre-functional	60	496	67	13	32	35	490	55	10	13	4	537	37	.	4
Beginner	3	–	–	–	–	2	–	–	–	–
Intermediate
Advanced
Full English Proficient
Title III First Year Exited
Title III Second+ Year Exited
English Speaker I	1	–	–	–	–	2	–	–	–	–	1	–	–	–	–

¹⁴ Note: Data marked '–' are suppressed because the subgroup contains fewer than 10 students.

	Elementary School (ES)					Middle School (MS)					High School (HS)				
	N	Scale Score		Ach. Level		N	Scale Score		Ach. Level		N	Scale Score		Ach. Level	
		Mean	SD	< 3	>= 3		Mean	SD	< 3	>= 3		Mean	SD	< 3	>= 3
English Speaker II	1430	505	62	191	792	1289	518	75	288	596	352	510	100	131	214
Pre-functional Waiver	1	—	—	—	—
Beginner Waiver	1	—	—	—	—
Intermediate Waiver
Advanced Waiver
ELIGIBLE FOR FREE OR REDUCED-PRICE LUNCH															
Full-Pay Meals	395	488	59	62	209	388	508	71	98	166	112	486	100	53	55
Free Meals	997	512	59	127	555	845	524	77	173	404	224	522	99	69	152
Reduced Meals	103	499	80	15	63	98	504	62	28	45	21	503	96	10	11
Unknown	1	—	—	—	—
IEP Disability Codes (Multiple Codes per Student)															
Severe Mental Disability	136	404	72	74	13	136	422	68	81	11	31	376	104	25	6
Moderate Mental Disability	333	491	48	45	181	404	502	48	108	178	133	496	71	56	73
Mild Mental Disability	389	540	41	9	267	365	567	57	14	230	98	581	77	6	90
Autism	400	499	44	57	220	302	504	68	93	122	70	472	90	39	30
Deaf/Blindness
Emotional Disability	17	547	37	.	12	10	—	—	—	—	1	—	—	—	—
Hearing Impaired	21	478	71	5	13	21	492	38	5	4	11	523	63	2	9
Learning Disability	44	565	48	.	35	27	616	57	.	21	3	—	—	—	—
Multiple-Disability	1	—	—	—	—	2	—	—	—	—	2	—	—	—	—
Other Health Impaired	79	509	72	8	43	63	550	63	6	39	14	550	105	5	9
Orthopedically Impaired	61	495	49	9	37	49	506	84	11	19	14	513	82	4	9
Speech or Language Impaired	1090	508	54	131	624	665	517	62	146	314	119	517	81	41	75
Traumatic Brain Injury	11	494	46	4	7	8	—	—	—	—	4	—	—	—	—
Visually Impaired	53	450	87	18	17	49	478	61	17	13	11	476	145	5	6
TOTAL	1496	505	62	204	827	1331	518	74	299	615	357	509	101	132	218

Exhibit I-4: Performance by Grade-Band Form and Student Age—Social Studies¹⁵

	Elementary School (ES)					Middle School (MS)				
	N	Scale Score		Ach. Level		N	Scale Score		Ach. Level	
		Mean	SD	< 3	>= 3		Mean	SD	< 3	>= 3
STUDENT'S AGE										
8	417	494	68	103	121
9	508	512	67	162	338
10	446	511	76	77	137	1	–	–	–	–
11	425	522	73	81	153
12	430	518	72	163	264
13	355	519	81	64	111
15										
16										
STUDENT'S ETHNICITY										
Unknown	1	–	–	–	–
Asian	22	485	44	8	8	11	484	61	4	7
African American	726	515	67	147	348	684	528	81	153	315
Hispanic	89	497	68	32	37	50	517	49	14	22
American Indian	4	–	–	–	–	3	–	–	–	–
Other	41	501	72	16	14	30	507	70	10	9
Hawaiian/Pacific Islander	2	–	–	–	–	1	–	–	–	–
White	611	503	70	164	246	552	516	70	150	236
STUDENT'S GENDER										
Unknown	1	–	–	–	–
Female	479	503	71	123	199	432	517	82	102	193
Male	1016	510	67	247	455	899	524	71	229	398
ESL (LANGUAGE)										
Unknown	1	–	–	–	–	2	–	–	–	–
Pre-functional	60	502	65	19	25	35	510	60	10	15
Beginner	3	–	–	–	–	2	–	–	–	–
Intermediate
Advanced
Full English Proficient
Title III First Year Exited
Title III Second+ Year Exited

¹⁵ Note: Data marked '–' are suppressed because the subgroup contains fewer than 10 students.

	Elementary School (ES)					Middle School (MS)				
	N	Scale Score		Ach. Level		N	Scale Score		Ach. Level	
		Mean	SD	< 3	>= 3		Mean	SD	< 3	>= 3
English Speaker I	1	–	–	–	–	2	–	–	–	–
English Speaker II	1430	508	69	350	626	1289	522	75	320	576
Pre-functional Waiver	1	–	–	–	–
Beginner Waiver	1	–	–	–	–
Intermediate Waiver
Advanced Waiver
ELIGIBLE FOR FREE OR REDUCED-PRICE LUNCH										
Full-Pay Meals	395	486	71	138	129	388	507	70	122	147
Free Meals	997	517	66	207	478	845	529	77	183	408
Reduced Meals	103	507	73	25	47	98	518	61	26	36
Unknown	1	–	–	–	–
IEP Disability Codes (Multiple Codes per Student)										
Severe Mental Disability	136	398	72	82	6	136	423	69	88	7
Moderate Mental Disability	333	488	60	112	118	404	509	53	116	153
Mild Mental Disability	389	550	43	24	258	365	569	55	21	242
Autism	400	501	53	124	146	302	506	58	95	109
Deaf/Blindness
Emotional Disability	17	541	37	1	5	10	–	–	–	–
Hearing Impaired	21	480	90	6	10	21	510	58	5	12
Learning Disability	44	562	46	2	26	27	629	73	1	18
Multiple-Disability	1	–	–	–	–	2	–	–	–	–
Other Health Impaired	79	521	70	16	45	63	539	71	11	34
Orthopedically Impaired	61	487	69	20	21	49	504	94	16	21
Speech or Language Impaired	1090	511	61	254	495	665	517	64	173	274
Traumatic Brain Injury	11	524	56	2	5	8	–	–	–	–
Visually Impaired	53	440	87	25	7	49	469	64	24	12
TOTAL	1496	508	69	370	654	1331	522	75	331	591